DOCUMENT RESUME

ED 227 440 CS 007 008

AUTHOR	Pearson, P. David, Ed.; Hansen, Jane, Ed.
TITLE	Reading: Theory, Research, and Practice. Twenty-Sixth Yearbook of the National Reading Conference.
INSTITUTION	National Reading Conference, Inc.
PUB DATE	77
NOTE	309p.; Proceedings of the annual meeting of the National Reading Conference (26th, Atlanta, GA, December 2-4, 1976).
PUB TYPE	Collected Works - Conference Proceedings (021) -- Reports - Research/Technical (143) -- Information Analyses (070)
EDRS PRICE	MF01/PC13 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; Higher Education; Language Processing; Learning Disabilities; *Learning Theories; Program Development; Reading Attitudes; *Reading Comprehension; *Reading Instruction; *Reading Processes; *Reading Programs; *Reading Research; Study Skills; Teacher Education
IDENTIFIERS	National Reading Conference (Organization)

ABSTRACT

The papers in this collection represent a wide spectrum of approaches, philosophies, and viewpoints of scholarly endeavor in their treatment of reading theories, research, and processes. The collection begins with an address by the president of the National Reading Conference concerning the tools for reading researchers. The remaining 53 articles are arranged according to the following categories: (1) reading and study skills, (2) reading and teacher education, (3) reading and program development, (4) reading and the affective domain, (5) reading and learning disability, (6) reading and word identification, (7) reading and comprehension, and (8) reading and language. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED227440

READING: THEORY, RESEARCH, and PRACTICE

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

TWENTY-SIXTH YEARBOOK OF THE NATIONAL READING CONFERENCE

Edited by
P. DAVID PEARSON
University of Minnesota

with the editorial assistance of
JANE HANSEN
University of Minnesota

Published by
The National Reading Conference, Inc.
Clemson, South Carolina 29631

1977

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
The National REading Conference, Inc.

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CS 007008

2

Copyright, 1977
THE NATIONAL READING CONFERENCE, INC.
All rights reserved. This book or parts thereof
may not be reproduced in any form without
permission of the publishers

Distributed by
THE NATIONAL READING CONFERENCE
Godfrey Hall
Clemson University
Clemson, South Carolina 29631

The ideas expressed in the papers published
in the yearbook are those of the authors and
do not necessarily represent the viewpoints
of the National Reading Conference

PRINTED IN THE UNITED STATES OF AMERICA
BY
MASON PUBLISHING COMPANY
St. Paul, Minnesota

OFFICERS AND BOARD OF DIRECTORS
OF THE
NATIONAL READING CONFERENCE, INC.

J. JAAP TUINMAN, *President*
Simon Fraser University

HARRY SINGER, *Vice-President*
University of California

GORDON GRAY, *Treasurer*
Clemson University

VONCILE C. MALLORY, *Secretary*
Florida Atlantic University

EDWARD FRY, *Past-President*
Rutgers University

FRANK GREENE, *Member*
McGill University

JOANNA WILLIAMS, *Member*
Columbia University

WAYNE OTTO, *Member*
University of Wisconsin

RON CARVER, *Member*
University of Missouri

RICHARD ROBINSON, *Convention Manager*
University of Missouri

P. DAVID PEARSON, *Yearbook Editor*
University of Minnesota

J. JAAP TUINMAN, *Editor, Journal of Reading Behavior*
Simon Fraser University

ANNE MARIE BERNAZZA HAASE
Publications Chairman
University of Arizona

EDITORIAL ADVISORY COMMITTEE
26th NRC YEARBOOK

DICK ALLINGTON
State University of New York at Albany

EUNICE ASKOV
Pennsylvania State University

MARK AULLS
McGill University

WILLIAM E. BLANTON
Appalachian State University

EMERY BLIESMER
Pennsylvania State University

RICHARD BLOOMER
University of Connecticut

DON CUNNINGHAM
Indiana University

BETH DAVEY
University of Maryland

PRISCILLA DRUM
University of California -
Santa Barbara

KEN DULIN
University of Wisconsin

JOHN GUTHRIE
International Reading Association

ANNE MARIE HAASE
University of Arizona

DOROTHY HANSEN
Brigham Young University

JEROME C. HARSTE
Indiana University

DALE JOHNSON
University of Wisconsin

AL KINGSTON
University of Georgia

MICHAEL L. KAMIL
Purdue University

RON LESLIE
New York University

MERVIN LYNCH
Northeastern University

MARTHA MAXWELL
University of California - Berkley

GEORGE McNINCH
University of Southern Mississippi

BONNIE J. F. MEYER
Arizona State University

BEVERLY MORRISON
University of Wisconsin

LEE MOUNTAIN
University of Houston

PHIL NACKE
University of Kentucky

ROBERT PALMATIER
Northwestern State University
of Louisiana

CHRISTOPHER RAMIG
Georgia State University

EARL F. RANKIN
University of Kentucky

ROBERT SCHREINER
University of Minnesota

HARRY SINGER
University of California - Riverside

ROBERT TIERNEY
University of Arizona

DAVID M. WARK
University of Minnesota

TABLE OF CONTENTS

Page

PRESIDENTIAL ADDRESS

Tools for Reading Researchers 1
Edward Fry

READING AND STUDY SKILLS

Effect of Student-generated Pre-questions and Post-statements on Immediate and Delayed Recall of Fourth Grade Social Studies Content: A Pilot Study 11
Laura Chodos, Sandra M. Gould, and Reuben R. Rusch

Effectiveness of a Study Management Course for "Non-traditional" Students 17
Kristina G. Hayward, Vincent P. Orlando, and Emery P. Bliesmer

Measuring and Evaluating Changes in Underlining and Note-Taking Skills 21
Marilyn M. Fairbanks and Carolyn Costello

The Effects of Experimenter Underlining and Expertise on the Retention of Textual Material 26
Steven M. [illegible], Diane H. [illegible], and David M. [illegible]

READING AND TEACHER EDUCATION

A New Hypothesis for Reading Teacher Research: Both the *Teaching* and *Learning* of Reading are Theoretically Based 32
Jerome C. Harste and Carolyn L. Burke

Developing and Evaluating a Criterion Referenced Test for Preservice Reading Education Students 41
Shirley B. Merlin, James L. Laffey, and Noel E. Bowling

The Relationship Between Teacher Assessment of Pupil Mastery and Standardized Reading Achievement Scores 46
Robert L. Aaron, Alfredo Stokes, and Etta Bates

Diagnostic Teaching in Content Area Reading 52
Mary M. Dupuis and Eunice N. Askov

Teacher Performance in Assessment of Comparative Reading Difficulty of Content Materials 60
Robert A. Palmatier and Susan S. Strader

Page

Competencies for College Developmental Reading Teachers 63
Rosalind Strichler and Joseph S. Nemeth

Stability of Teacher Effect on Pupils' Reading Achievement over a Two Year Period and Its Relation to Instructional Emphases 69
William H. Rupley

An Analysis of Blacks in American History Textbooks as Perceived by Social Studies Teachers 73
Ann B. Dolgin

Reading Teachers' Knowledge of and Attitudes toward Black English 76
Kathleen Stumpf Jongsma

READING AND PROGRAM DEVELOPMENT

The Use of Receptive and Expressive Language Development Activities for Improving Students' Reading Achievement 82
Hal W. Seaton and Sharon D. Smith

Twelve Year Study of Pupils Who Were Underachieving in Reading in First Grade 88
Anabel P. Newman

Sequence Strategies for Teaching Reading Comprehension with the Cloze Procedure 92
Earl F. Rankin

Developing an Essential Reading Skills Program for Post-Elementary Students 99
Roger Klumb and Wayne Otto

Improving High School/College Articulation for Compensatory Students Through Early Basic Skills Evaluation 105
Joyce D. Johnston

Impact of an Experimental Reading-Study Skills Course on High-Risk Student Success in a Community College 110
Eleanor Haburton

Evaluating the College Reading Program: A Management Audit 115
Ann Marie Bernazza Haase, Shirley DeShields, and Mervin D. Lynch

Page

READING AND THE AFFECTIVE DOMAIN

Growth in Reading as a Correlate of Pupil Classroom Behavior 118
Homer Coker and Jeffrey L. Lorentz

Effect of Uninterrupted Sustained Silent Reading and of Reading Games on Changes in Secondary Student's Reading Attitudes 126
Larry Mikulecky and Anne Wolf

Predicting Reading Performance at the Secondary Level through the Utilization of a Student Self-Assessment Inventory................. 131
Charles W. Peters

Masculine, Feminine, and Androgynous College-Age Readers' Story-Plot Preferences Related to Gender of Protagonist................. 136
Patricia L. Anders and Ken L. Dulin

A Preliminary Study of the Relationships Among Animal Identifications, Learning Preferences, Reading, and Reading Related Abilities... 139
Anthony V. Manzo

Issues of Sexism and Ethnicity in Reading.............................. 147
Albert J. Kingston, Terry L. Lovelace, and Molly M. Wilson

READING AND LEARNING DISABILITY

The Myths and Realities of Severe Reading and Related Learning Disabilities... 150
Jules C. Abrams

A Multidisciplinary Approach to the Etiology and Management of Learning Disability 158
Sheila K. Hollander

Special Considerations in Teaching LD Children to Read........... 163
Catherine Morsink and Wayne Otto

READING AND WORD IDENTIFICATION

Investigating the Role of Meaning in Mediated Word Identification . 168
Patricia M. Cunningham

Functional Word Identification Ability of Poor Readers 172
Martha C. Evans, Nancy E. Taylor, and Irene H. Blum

Word Boundaries Revisited: A First Grade Study.................. 179
Josephine S. Goldsmith and Mark J. Nicolich

Page

The Effects of Grade Level and School Setting on the Development of Sensitivity to Orthographic Structure 183
Jerome A. Niles, Aileen Grunder, and Carol Wimmer

Developmental Trends in the Discrimination of High Frequency Words 187
Richard L. Allington

What Level of Representation do Children Expect to Find in Print? 193
Charles Read

READING AND COMPREHENSION

Psycholinguistic Analysis of Cloze Responses 200
Joseph L. Vaughan, Jr., Robert J. Tierney, and Madelon Alpert

Clause Strategies as a Factor in Reading Comprehension 203
R. Scott Baldwin

Subvocalization — Asset, Liability, or Not? 209
Carl T. Cline, Jr.

Organization in Prose and Memory: Research with Application to Reading Comprehension 214
Bonnie J. F. Meyer

A Discourse Structure Analysis of the Comprehension of Rapid Readers 221
Dorothy M. Hansen

Reading Ability and Memory for Sentences and Words 227
James T. Fleming and Fred W. Ohnmacht

The Effects of Non-Prose Textual Devices on Retention of Major Concepts and Supporting Details with and without Specific Instructions in Their Use 232
Jack N. Cole

Childrens' Retrospective Oral Responses to Silent Reading Test Items 236
E. Noel Bowling and James L. Laffey

An Investigation of the Relationships of Selected Components of Readability and Comprehension at the Secondary School Level 244
Gary Lee Shaffer

Page

Readability and Other Factors in College Reading Tests: A Critique of the Diagnostic Reading Test, The Nelson-Denny Reading Test, and the McGraw-Hill Basic Reading Test 253
Chester E. Tillman

The Raygor Readability Estimate: A Quick and Easy Way to Determine Difficulty 259
Alton L. Raygor

READING AND LANGUAGE

Paradigmatic/Syntagmatic Behavior and the Reading Process 264
James A. Dinnan

The Effects of Paradigmatic Language Training on the Reading Process: An Experimental Study 267
Lora M. Cartelli

Paradigmatic-Syntagmatic Associations, Reading Achievement and Dominance in Culturally Disadvantaged Secondary Students 272
Albert J. Kingston, James A. Dinnan, and Karlyne King

Paradigmatic-Syntagmatic Responses of Bilingual Students and Reading Comprehension 275
Paul A. Lamarre

A Preliminary Investigation of Oral Language of Bilingual Children and the Beginning Reading Process 280
Sally Kaminsky, Rose Harrison, and Martha J. Bell

The Effect of Lexically Ambiguous Nouns Embedded in a Reading Task for Children 284
Diane Surle Stupay

Children's Understanding of Linguistic Connectives: A Review of Selected Literature and Implications for Reading Research 292
Sean A. Walmsley

FOREWORD

The papers in this volume represent a wide range of topics, concerns, and scholarly genres. There are empirical studies, state of the art position papers, reviews of literature, evaluation reports, program development reports, and think pieces. They share two common threads. First, they all address issues about reading; second, they were all selected for publication with a single criterion in mind: good scholarship deserves a home.

The work of the editorial advisory board deserves special note. All its members worked carefully and swiftly to assist me in making decisions about the appropriateness of articles submitted for review.

Many students and staff here at the University of Minnesota deserve a note of appreciation for their work and support. In particular, Beth Marr, Ginny Allery, and Lorri Neilsen helped to set up an elaborate procedure for the flow of manuscripts to and from reviewers, authors and printers. Diana Pinkston made the system work. Josephine Zimmer and her secretarial pool kindly retyped manuscripts when necessary. Jane Hansen aided me in the copy editing process.

I hope readers find the papers useful and interesting. They came out of a useful and interesting conference. There is a certain sadness associated with their publication. They have been around my office so long, I sort of hate to see them leaving to enter the world of scholarship. Alas! Even penguins leave their nest.

P. David Pearson
University of Minnesota

EDWARD FRY
Rutgers University

PRESIDENTIAL ADDRESS

TOOLS FOR READING RESEARCHERS

Reading research is in an epistemological crises; we don't know what we don't know. Researchers are in a situation not unlike the allegorical blind men in India. Upon discovering an elephant one grabbed ahold of his tail and maintained the elephant was like a rope, another put his arms around a leg and maintained the elephant was like a tree, while a third ran a hand along his side and assured his fellows that the elephant resembled a wall.

Among reading researchers and model builders we have the rope group, the tree trunk schools, and the wall advocates. Staunchly clinging to their own data, they have developed even finer methods of measuring the rope's index of flexibility, the wall's height, and the tree trunk's mean diameter, based upon randomly selected samples according to a theoretical model, refined by analysis of variance, and including a standard error of measurement significant to $p <$ .001. Broad-minded, interdisciplinary researchers have correlated leg diameter and tail diameter and found them to be .73.

This paper is a plea for a little more data before we build an elephant; it is a request that we stop building models until we get more data. I am aware of the argument that models generate and synthesize isolated research studies, and I am also aware that by asking model builders to stop, I am not going to change their behavior one whit. Man's need for closure or perhaps his ego drive for achieving *the* $E = MC^2$ of reading is too strong. However, perhaps I can encourage those who are rightly skeptical of existing models, those who seek hard data from a variety of sources, and those who try to piece together smaller bits of data into firm concepts, to open up the system into wider areas with an increasing refinement of techniques.

I might ask the model builders to look at the recent history of human learning in psychology. What ever happened to the "laws of learning" that we used to teach a few decades ago. They were built on too narrow a data base and toppled over in the light of new information.

With that concept in mind I will discuss some research tools at hand. These tools are not totally new, but by reviewing some of them, it might stimulate thinking for alternate forms of knowledge generation. Most of us wear several hats: we disseminate knowledge about reading through teaching and writing, we are practitioners trying to improve reading acquisition skills, and we either produce or are closely related to, the production and interpretation of new knowledge by research.

HARDWARE TOOLS

Eye Movement

Eye movement research is certainly not new. Javel started using it around in 1879, and Tinker almost exhausted its use in the 1930's and 1940's. But it clearly fits my definition of a tool, and it still has relevance. It is a nice tool. It can easily be observed by anyone (and hence has a lot of surface validity), and it can be objectified, quantified, refined, and subjected to powerful computer analysis.

On the Javel level you can have a class of undergraduates pair themselves off: one student reads and the other student watches the eye movements of the reader to bring meaning to terms like fixation, saccadic movement, return sweep, and regression. On the Tinker level you can purchase eye movement recording instruments which give you graphic records of eye movements and enable you to generate data showing such things as a growth pattern between first grade and college in decreasing duration of fixation, decreasing number of regressions, and increased span of recognition.

Techniques for observing eye movements have progressed from lying a mirror alongside of a book so someone standing behind the reader can observe eye movement, to electro-oculogram devices which can sense eye movement by treating the eye ball as a very delicate flashlight battery which emits an electrical potential. Shifts in the + /- pole can be sensed by placing an electrode on the nose and forehead. These and improved camera devices can study eye movement in greater detail permitting such insights as Rayner's (1975) observation about rather long fixations on infrequent words such as "rendezvous" and "cache". Thus, we can see that duration of fixation is perhaps related to word difficulty. Since computers can be programmed to summate fixation durations, and since, through other research we know that word difficulty is related to readability, it can be seen that readability research in the near future could be enhanced by an electro-oculogram plus a small computer. Since readability is related to, or in some aspects is the obverse side of comprehension, we can see how eye movement might be a measure of comprehension.

In traditional comprehension research and evaluation, questions following the reading of a passage are used, but this raises a problem: "Is the student failing to comprehend the passage or failing to comprehend the question?" The cloze technique avoids this problem. Another way of avoiding it, however, might be with eye movement or *fixation duration* scores. This is what I mean by using research tools — developing new tools and the facility to use them, to pry further into such simple and central problems as "What is involved in comprehension?" or "Why is passage A harder to read than passage B?" Reading efficiency or rate might better be improved or at least studied by looking at eye movements.

Perception

The perception of letters, words, phrases and larger segments has occupied a great deal of research time. As a series of tools, I find it rather boring; it has been around so long and parts of it seem so little changed. In 1908 Huey discussed a study by Messmer in which he found that the longer letters which projected above the line are usually the dominating ones. In 1968 Gibson and her co-workers were interested in distinctive features and:

> employed a same-different judgment of two letters exposed simultaneously by projecting on a small screen. If the subject thought they were identical, he pressed one button; if he thought they were different, he pressed another. His latency, the time he took to respond, was the index used as well as the few errors that were made.
>
> (Gibson and Levin, 1975)

Huey and other turn-of-the-century researchers were also interested in deleting parts of words. They found, for example, that the top half deletion made reading harder than bottom half deletion and that the first half of a word was more important than the second half. Deletions can also be accomplished by a variety of techniques, such as placing various screens over words or parts of passages. They offer us another technique of determining word difficulty or familiarity.

The tachistoscope, that overworked workhouse of verbal learning research, has enjoyed an amazing longevity. Dodge was using it at the turn of the century, and it is still a staple in experiments that require a short exposure and standard viewing conditions.

Some modern techniques of looking at perception related to the reading task involve embedding words or phrases in different matricies or in running words together to see if the subject can detect the word boundaries (Klein and Klein, 1972).

Eye Voice Span

The span between where the voice is and where the eye is when reading aloud can be easily observed by simply sliding a card over the page while the subject is reading aloud and counting the number of words he continues saying. In 1897 Quartz noted that the eye voice span was longer at the beginning of a line than toward the end with averages varying from 7.4 words at the beginning to 3.8 words at the end of a line. He also noted that the span decreased to zero when an unfamiliar word was encountered. Somewhat more interesting, he noted a close correlation between the increase of eye-voice span and an increase in rate of reading. He felt that "a considerable distance between eye and voice is a condition of intelligent and intelligible reading" (Huey, 1908).

Eye voice span research is still with us. Geyer (1968) looked at it in terms of elapse time rather than number of words, but modern laboratory equipment might be able to provide a whole new range of eye voice span data by continuously linking eye fixation point with voice. Perhaps greater refinements can link eye movement to subvocalization.

Vocalization

There is a lot of overlap between speech research and reading research and many of the tools are the same. To a large extent we share a common history. Reading researchers have long been interested in subvocalization (or inner speech) while reading. It can be reported subjectively by nearly everyone, but trying to objectify it is difficult. In 1900 Curtis studied covert oral activity by mechanically sensing larynx movement; modern researchers tend to use an electro-myograph or (EMG), a recording device for minute electrical changes associated with muscle activation. Sensors are placed on the throat or mouth areas and slight electrical changes can be noted and associated with silent reading activity long before muscle movement can be detected.

Hardyck (1968) used an EMG as a biofeedback device to lower subvocalization. This lowered EMG did not result in rate increase, but did result in reports of less fatigue with reading for periods of one to three hours. The lowered fatigue seems to be related to breathing pattern; a higher subvocalizer breathes more like a person talking (Hardyck and Petrenovich, 1969). However, lowering subvocalization may lower comprehension in poor readers.

An interesting tool for the study of speech is the voice spectrograph. This device, which is also called the sound or speech spectrograph, is essentially a visual display of speech or other sounds with time moving along the horizontal axis, pitch along the vertical axis, and intensity by degrees of darkness (Prestigiacomo, 1957). Words can literally be shown graphically. Among other interesting things, it shows no temporal pause between phonemes or, in many instances, between words. Spoken speech is a continuous flow. It calls into question what we are really talking about when we talk about a "word." In terms of sound production, a word is not a unit. Perhaps it is a writing convention. Perhaps it derives its definition from a cognitive base. At least, if you are interested in demonstrating what a word really is, you will get precious little help from a spectrograph, and this little bit of data could force some changes in traditional thinking about the reading process, or perhaps even something practical like, how we set about teaching children to read. If words are not inherent in speech, it could be that words are not the most efficient unit of reading instruction.

Electroencephology

If we are interested in minute muscle changes associated with reading, the next logical area of progress is to study brain activity. EEG's, or electroencephalograms, for all their ability to detect gross activity like epileptic seizures, or generalized states of consciousness like sleep, seem to be a bit too crude for specific reading research. But knowing about this reminds us that a good bit of psycho-physiological research is ultimately related to thinking and thinking to reading. Detecting electrical activity associated with the brain was done with animals in the 1870's and external detection of brain waves was started in the 1920's. Today it is a very active field. Reading researchers have most recently shown an interest in it seeking an illusive phantom called dyslexia in an attempt to explain reading failure. The relationship between reading failure and brain wave irregularities has not been shown in most instances. However, this does not mean that in the future neurological research does not have promise for understanding both reading failure and more importantly, the normal reading process.

Heart Rate

Wark (1971) reported results of a small but unique study utilizing a digital cardiotachometer, a relatively simple, off-the-shelf medical device, that almost instantly reports heart rate from electrodes attached to the arms. A base heart rate is obtained by having the subjects rest with their eyes closed. When asked to read, heart rate decreased. When they were asked to answer questions, the heart rate went above base rate. Wark concluded, "heart rate seems to be implicated in the reading process in a way not previously suspected."

Should continued investigations show relationships between heart rate and aspects of reading it is possible that some reading clinicians may be interested in affecting reading performance by biofeedback training. Physiological investigations have shown a remarkable ability on the part of yogis to reduce heart rate from 63 to 24 beats per minute. Some biofeedback training has worked but has shown less dramatic ability to affect heart rate (Brown, 1972).

Biofeedback

A number of the devices that we have just been discussing such as cardiotachometers, EEG and EMG have been used in the recent biofeedback movement. Biofeedback is the process of enhancing information about body functions so that they can be put under conscious control or conditioned or trained to more desirable functioning. Early enthusiasm and promise have been somewhat dampened by inexactitude, the discovery that often cheap instruments are unsatisfactory, and the unpleasant fact that much of it is plain hard work. The results are less than miraculous, but I doubt that it will disappear from either medical or learning research. We are prisoners of our own bodies and are at times dominated by our unconscious. The desire for greater freedom and the perennial intrigue of the mind-body problem is not too distant from the subject of this paper: epistemology.

SOFTWARE TOOLS

A research tool need not be a piece of hardware employing electronics; a good paper and pencil test can be every bit as important as a heart beat counter or a cardiotachometer.

Ability Tests

In reading, researchers and practitioners often seek to find either special or generalized abilities that predict reading achievement. Predicting future ability to read often involves some type of perceptual discrimination. I welcome this in the research area, but I am horrified at the erroneous judgments made by practitioners, such as first grade teachers who withhold reading instruction based on a readiness instrument that only correlates with success at about .50. We must all be cautious when training teachers not to confuse research instruments with those that should be used in the classroom.

For the moment, let's clearly separate measuring achievement from measuring ability. In measuring abilities, it is surprising to me that there hasn't been more work in a special ability called "reading." For many years we have had tests of special abilities in music, special abilities in mathematics, and other fields, yet we almost seem to ignore the possibility that reading ability might be special or might follow the normal distribution curve. We seem to accept the school boards' assumption that everybody should read at the same level or at least that the normal distribution curve shouldn't apply below the mean.

Achievement Tests

Thorndike and Hagen (1969) state that future historians will probably call the current period, the last twenty years, the period of the increased utilization of the achievement battery. Some type of reading achievement test is used at least several times in most student school careers in most school districts. Even with the millions of tests being given, we are almost at a primitive stage in content, quality. By simply looking at several major reading achievement batteries, it is readily apparent that there is little agreement on content or types of skills tested. Davis (1971) attempted to statistically analyze reading comprehension items and had great difficulty separating many factors; vocabulary stood out from all the rest, but even the terminology used to describe the factors is somewhat ambiguous. At most, he had eight categories, but some of those were barely independent.

The current increased use of test management systems that incorporate criterion referenced tests into a prescriptive system for classrooms should tend to heighten our concern for factors that might or might not be relevant to measuring reading success, and to factors that should be included into the teaching of reading. I recently had an interesting argument with an undergraduate class that insisted that the reading field knew what factors should be tested and taught. I couldn't convince them that we were so dumb.

The strength of the criterion test movement is that they measure skills to an absolute mastery criterion, independent of a norm reference group. Their weakness is that there is not strong research basis for determining what skills to measure.

An interesting blend of achievement and criterion referenced tests can be seen in the Basic Word Vocabulary Test completed by Dupuy for the National Center for Health Statistics (1975). Dupuy is a trained statistician, but being unacquainted with the reading field, he boldly set out to measure the vocabulary size of the non-institutional population of the United States. After overcoming initial hurdles, he succeeded quite well. The first hurdle was to find out, "What is a 'word'?" If, for example, we look at "run, runs, running, and ran," are they really one word, two words, or four words? Carroll in the American Heritage study would call them four words. A second hurdle is the so-called semantic count; is "run" to the store the same "run" as a "run" on the stock market?

Dupuy solved the problem by going to the unabridged dictionary and taking a 1% sample of main entries, omitting derivations, foreign, technical and compound words, then testing only the first meaning. Despite the fact that Carroll found 87,000 words and unabridged dictionaries claim over 200,000 entries, Dupuy found only 12,300 basic words. He made a test of 1% or 123 items and administered it to over 3,000 Virginia students who also had taken other normed vocabulary tests. He was thus able to anchor his test to U.S. norms and have a criterion score which was the total number of basic words known by multiplying raw score by 100. One interesting interpretation of his findings is that fifth graders can read and know the meaning of about one-quarter of the basic words, tenth graders know about half the words, and Ph.D.'s know about three-quarters.

An interesting implication of this is the rather solid evidence that half the vocabulary growth takes place between grades 5 and 10. Some of us have been so nearly brainwashed about the importance of language development in early childhood, that this information might give us some heart about what is being done, and what needs to be done, in the middle school years. Last, but not least, Dupuy's test is interesting in its extreme range; few other instruments can measure growth in any reading skill from third grade to beyond Ph.D. with both norms and a clearly defined criterion. This tool has implications for both longitudinal school achievement and language development in a neglected range. It also gives a Vocabulary Development Quotient which is similar to an Intelligence Quotient.

Before we leave the achievement tests, let me suggest that I suspect that there is some valuable data hidden in the computer memories of most major test companies. A majority of schools now pay for computer scoring, and they program pages of output on individual, class, school and district score reports. They have data on every item, age of children, sex, school, sub-test scores in other subjects, IQ, time of year, and often repeated testing in later years. Studies could investigate many possibilities. To cite a few of the many possible examples:

1. maybe a few math skills can predict vocabulary growth years later;
2. maybe certain reading items indicate future failure or underachievement;
3. maybe there are regional differences in skill development;
4. maybe some reading skills don't make a bit of difference to mature comprehension;
5. maybe some types of items favor boys or girls.

Research Design

Finally, I would like to suggest that research design and techniques are very definitely tools. Controlling for *training time* or *order of presentation* can be every bit as important in trying to gain knowledge as using the right instrument.

Subjective judgment is not dead. Oftentimes, it is the only way to get at attitude or reading ease. Singer (1975) successfully had college students judge readability by comparing unknown passages with passages of known difficulty (the SEER Technique). Carver (1976) utilized carefully selected and trained judges to compare test passages against standard passages in the Rauding Scale, and at Rutgers, we have been attempting to verify the Kernel Distance Theory by having students judge sentence pairs of equal vocabulary and length, but differing syntax.

Our investigations should vary in size. Much information about brain functioning has and will continue to be gained by studying individual aphasia victims, or looking at the longitudinal academic growth of a few children. Clinical studies, especially, make good use of controlled or structured observations which are partly exemplified by the use of a checklist.

Large scale studies can also give us information that is not obtainable by any other means. The National Assessment of Educational Progress is valuable and should be expanded. We need more data on the overall reading ability of the nation.

The USOE First Grade Studies in the mid-1960's were the largest scale reading experiment ever. They helped us to see what was going on in different instruction methods, and they slowed down the claims of zealots who had some sure-cure methods. Something like the First Grade studies should be repeated about every decade. Right now we need large-scale comparisons of test management systems like the Wisconsin Design with book controlled systems like the basal readers, and studies of individualization versus traditional grouping.

The First Grade Studies also raised an interesting question, yet unanswered. Why did a reading method like the Language Experience Approach, which has almost no specific skills emphasis, succeed in teaching reading as well as other highly skills-oriented programs like the so-called linguistic or phonics approach?

Large-scale studies might not be important for model builders, but they are certainly important for superintendents, curriculum committees, and classroom teachers trying to decide what to do next year.

They also have important implications for society as a whole. For example, one important finding of the National Assessment that confirmed smaller studies, was the high correlation between reading achievement and socio-economic status. Our newspapers which castigate reading methods usually fail to point out the solid correlation between poor reading achievement and such SES factors as parent education, number of parents in the home, or family income.

Miscellany

The computer is too big to be called a tool; it is a whole garage full of tools. It has a myriad of uses besides its standard uses for large-scale data handling and high level statistical analyses. Computer simulation has been an influence on the way we analyze tasks and try to reproduce them. MIT is now closing in on a computer that can read orally from any printed page. Computer control of the teaching process as in CAI offer not just a way to supplement instruction which is good, but a way of controlling teaching variables so that one variable can be changed while all others are held constant. On-line computer monitoring of sensitive instrumentation promises to discover things that the human eye or the human mind cannot detect. Its use in information search and retrieval are still in an infancy period.

A list of research tools would not be complete without mentioning research summaries and document collections. Those who don't know history are condemned to repeat it. The *NRC Yearbook* summaries by Bleismer and others and the *Reading Research Quarterly* annual summaries are very useful to current and future researchers. Incidentally, the total *William S. Gray Collection of Reading* documents has just become available on microfiche from the Hofstra University. The ERIC collection is another valuable resource tool, as are Kling's (1971) literature summaries and Corder's (1971) summaries of reading research for the USOE Targeted Research. NRC members might be interested to learn that our first eight yearbooks, most of which are out-of-print, are now available on microfiche from ERIC. Our old papers have received a kind of technological posterity.

We face an exciting future in reading research. Undoubtedly that research will uncover new tools. In the meantime, we already possess a number of useful tools — tools that properly refined and sensibly used will serve us well in uncovering new secrets, verifying old intuitions, and following good hunches about this remarkable process called reading.

Finale

And now the time has come for you to see what research tools have wrought. This year's program at the National Reading Conference is one of the very best. Jaap Tuinman, your President for next year, has labored long and hard to put it together, and most of you have spent countless hours at the typewriter, at the computer center, and chasing down subjects and references to put together your presentations. Most of the research reported here is not funded, or rather, it is funded by the sweat and overtime of the individual investigator. Even when it is funded, there is usually a measure of extra effort which can only be ruled by professional interest, not dollars. So, for all those long hours you have put in to improve the reading of children and adults in America, I for one, would like to say, "Thank you very much."

REFERENCES

Bio-Feedback Systems (catalogue). Boulder, Colorado, 1975.

Brown, B. *New mind, new body: Biofeedback.* New York: Harper Row, 1972.

Carver, R. Measuring prose difficulty using the rauding scale. *Reading Research Quarterly,* 1975-76, *11,* 660-685.

Corder, R. *The information base for reading: A critical review of the information base for current assumptions regarding the status of instruction and achievement in reading in the United States.* Educational Testing Service, Berkeley, California, 1971. (ERIC Document Reproduction Service No. ED 054 922)

Davis, F. B. Psychometric research on comprehension in reading. In M. King (Principal Investigator), *The literature of research in reading with emphasis on models.* New Brunswick, N.J.: Rutgers University, 1971.

Dupuy, H. J. *Basic word vocabulary test.* Highland Park, N.J.: Dreier Educational Systems, 1975.

Geyer, J. J. Perceptual systems in reading: The prediction of a temporal eye-voice span constant. In H. K. Smith (Ed.), *Perception and reading.* Newark, Del.: International Reading Association, 1968.

Gibson, E. J. & Levin, H. *The psychology of reading.* Cambridge, Mass.: The MIT Press, 1975.

Hardyck, C. D. & Petrinovich, L. F. Treatment of sub-vocal speech during reading, *Journal of Reading.* 1969, *12,* 361-368.

Hardyck, C. D. Petrinovich, L. F., & Ellsworth, D. W. Feedback of speech muscle activity during silent reading: Rapid extinction. *Science,* 1966, *159,* 1467-1468.

Huey, E. B. *The psychology and pedagogy of reading.* Cambridge, Mass.: MIT Press, 1968. (Originally published, 1908.)

Klein, G. A. & Klein, H. A. Word identification as a function of contextual information. *American Journal of Psychology,* 1973, *82,* 399-406.

Kling, M. (Principal Investigator), Davis, F. (Ed.). *The literature of research on reading with emphasis on models.* New Brunswick, N.J.: Rutgers University, 1971. (Available from ERIC on microfiche or in soft cover book from IRIS Corp., Box 372, East Brunswick, N.J. 08816.)

McGuigan, F. J., Keller, B., & Stanton, E. Covert language responses during silent reading. *Journal of Educational Psychology,* 1964, *55,* 339-343.

Prestigiacomo, A. J. Plastic tape sound spectrograph. *Journal of Speech and Hearing,* 1957, *22,* 321-327.

Raynor, K. *Eye movements in reading: Historical trends and current developments.* Unpublished Manuscript, University of Rochester, 1975.

Singer, H. "The SEER Technique: A non-computational procedure for quickly estimating readability level." *Journal of Reading Behavior,* 1975, *7,* 254-267.

Sound Spectrograph, Series 700. Somerville, N.J.: Voice Identification Laboratories.

Thorndike, R. L. & Hagen, E. *Measurement and evaluation in psychology and education.* New York: John Wiley and Sons, 1969.

Tinker, M. The study of eye movements in reading. *Psychological Bulletin*, 1958, *55*, 125-232.

Wark, D. M., Tostenrud, D., & Nelms, G. Heart rate and reading. *Yearbook, National Reading Conference*, Clemson, N.C., 1972.

William S. Gray Research Collection (9,000 documents and 18,000 index cards; 300,000 pages). Microfiche Edition, Alvina Treut Burrows Institute, Hofstra University, Hempstead, New York, 1976.

LAURA B. CHODOS
SANDRA M. GOULD
REUBEN R. RUSCH
State University of New York at Albany, NY

EFFECT OF STUDENT-GENERATED PRE-QUESTIONS AND POST-STATEMENTS ON IMMEDIATE AND DELAYED RECALL OF FOURTH GRADE SOCIAL STUDIES CONTENT

The absence of effective organizational strategies has been suggested as a possible cause for some comprehension difficulties. It is generally assumed that the greater the degree of organization which can be imposed on material to be learned, the better the recall. Wiener and Cromer (1967) suggested that students with ineffective organizational strategies could be assisted in either of two ways: 1) alter the organization of the material; or 2) train the students to adopt the organizational strategies used by good readers.

Considerable research is now available in support of a facilitative effect from various types of text alteration (Frase, 1970; Ausubel, 1960; Rothkopf and Kaplan, 1972; Rickards, 1976). Little is known however, about whether such approaches have any lasting impact on the student's comprehension ability when the imposed organizational strategies are removed. In fact, with few exceptions (Natkin and Stahler, 1969; Peeck, 1970; Boker, 1975), researchers in the area of "mathemagenic" behaviors (Rothkopf, 1965) have been concerned only with immediate retention measures. It is essential that future studies be concerned with the long term effects of any procedure believed to stimulate student behavior leading to learning; i.e., mathemagenic behavior.

Much less research has been done to explore the second alternative for assisting students, that is, training the reader to apply more effective organizational strategies himself (Allington, 1975). Levin (1972) found that instruction to use imagery led to improved comprehension among fourth grade readers. Frase and Schwartz (1975) found that recall of high school students and college freshmen was increased, compared to a study only condition, when students were told to engage in question production. Duffy (1973) suggested a type of student-generated questioning focused on teacher-selected key words. Schwartz and Scheff (1975) described a modification of the Directed Reading Thinking Activity (Stauffer, 1964) in which students are involved in questioning.

The use of student-generated questions as a mathemagenic has both theoretical and practical appeal. Its use is most directly traceable to Frank Smith (1975) who defines comprehension as the condition of having one's cognitive questions answered. If we can accept the possibility that some readers do not comprehend because either they did not ask or asked the wrong questions during reading, then a procedure which encourages student-initiated questions is supportable. However, the few studies which examined student-generated questioning as an organizational strategy have produced ambiguous results. Morse (1975) found no facilitative effects from experimenter *or* reader generation of questions, although ceiling effects may have masked potential differences. Smith (1972), in a six week experiment, studied the effectiveness of training students to generate questions. Primarily interested in quantifying the number of questions students generated after training, Smith found no significant differences in post-training scores on the Gates-MacGinitie comprehension sub-test; this is not surprising since it is generally accepted that standardized tests such as the Gates-MacGinitie are not sensitive to short term increment gain.

Keeping this research in mind, we initiated the present pilot study. Deliberately choosing heterogeneous fourth grade classes, we asked the question: Does instructing students to a) generate questions before reading; b) to write summary statements after reading, or c) doing a combination of a and b have any observable effect on immediate and delayed recall of social studies content? We were particularly interested in the delayed recall measure, being influenced in part by the provocative findings in studies such as Berlyne (1966) and Natkin and Stahler (1969), which suggest that questions may be used to generate epistemic curiosity and that the sparse spacing of questions may have positive effects on long term retention. According to Natkin and Stahler, such procedures might allow both arousal and mathemagenic effects to emerge (1969, p. 431).

We recognize that more important questions cannot be answered by this pilot study. For example, in our efforts to control time for all treatments, our tasks had to be carried out over two days. Because of these limits, there was no way to test study/only condition over the whole passage at one sitting. Consequently, we could not assess the mathemagenic effect of interspersed testing adequately. In addition, our decision to use recall of specific facts as the dependent variable was influenced by our desire to construct passage dependent questions; i.e., those questions which cannot be answered without reading the passage (Tuinman, 1973, 1974; Allington et. al., 1977). This decision excluded discussion of the possible effects of student-generated behavior on more complex cognitive processing.

METHOD

Materials and Procedure

A passage of about 1600 words was selected from a fourth grade social studies text and divided into four equivalent parts of about 400 words each. Each part represented a natural division in the events of the story. The story was Jacob Riis: To Help the Poor, Exploring with American Heroes, R. A. Brown, Follett Publishers, 1976. The text was not in use in the three schools in this study.

Experimenters developed 20 recall questions for each of the four parts of the story and randomly assigned ten of these to Form A for immediate recall and ten to Form B for delayed recall one week later. A split-half procedure was used in the delayed recall test only by randomly assigning students in each class to an odd-numbered set or an even-numbered set of 20 questions. The question sets were ordered so that each class would have an equal number of odd and even-numbered question sets. The order of question presentation for both the immediate and the delayed recall tests followed the event as they occurred in the story.

Scoring

The 9600 items in all the recall tests were corrected by two experimenters according to a predetermined criteria of acceptable responses. Both the acceptable responses in the criteria and the answers given were checked with each experimenter working separately and then working together to check each other's decisions.

The criteria employed in the determination of acceptable answers are reminiscent of the categories developed in research by Pearson and Nicholson (1976). Acceptable answers were either "textually explicit" or "textually implicit." Responses which could have been "guessed" (script intrusion) or incorrect inferences from the text (text intrusion) were not accepted. An example given below illustrates the decision process.

Ex. 1: The children were playing ________________ on the grass.

Leap frog	textually explicit	— acceptable
games	textually implicit	- acceptable
all over the	script intrusion	- not acceptable
happily	script intrusion	- not acceptable
cowboys	text intrusion	- not acceptable

Design

Six heterogeneously grouped fourth grade classrooms were randomly assigned six different treatments. Two interior measures were used. A post-test was administered consisting of four parts of ten items each. One week later, a similar retention test was administered consisting of four parts of *five* items each.

Description of the Treatments

Treatment #1 Student-generated questioning from a summary, a pre-reading activity.
A summary of about 50 words was read silently, and students were instructed to "write four questions in the spaces provided that you might expect to find answers for from reading the story that follows." Students were instructed also "don't worry if you cannot think of four questions. Do the best that you can." Time - 5 minutes. Students read the passage. Time - 5 minutes. Students answered ten recall questions. Time - 6 minutes. This procedure was repeated for all four parts of the story.

Treatment #2 Student-generated statements following reading.
Students read the passage. Time - 5 minutes. Following this they wrote four statements in the spaces provided. Time - 5 minutes. Students were instructed to "do the best that you can. Don't worry if you cannot think of four statements." Students answered ten recall questions. Time - 6 minutes. This procedure was repeated for all four parts of the story.

Treatment #3 Student-generated questions from a summary, and student-generated statements following reading.
Students followed procedures outlined for treatments #1 and #2. This treatment required 5 additional minutes in all. This procedure was repeated for all four parts of the story.

Treatment #4 Read/study with summary paragraph as an advance organizer.
With the summary as an advance organizer, students read/studied the story with instructions that "you will have plenty of time to read, reread and/or study the story." Time: 10 minutes. Students answered ten recall questions. Time: 6 minutes. This procedure was repeated for all four parts of the story.

Treatment #5 Read/study only.
Students read/studied story. Time: 10 minutes. Students answered ten recall questions. Time: 6 minutes. This procedure was repeated for all four parts of the story. Instructions included, "you will have plenty of time to read, reread or study the story."

Treatment #6 Read/study two parts at a time. 20 recall questions at the end.
Students read/studied two parts of the story at a time. Time: 20 minutes. Students answered 20 recall questions. Time: 12 minutes. Instructions included "you will have plenty of time to read/reread or study this part of the story." This procedure was repeated once again for parts 3 and 4.

RESULTS

The descriptive statistics pertinent to the results of the study are found in Table 1. The post-test is defined as the sum of the four subtests administered after each reading passage. The total possible score is 40, ten on each part.

The retention test, an alternate form of the same test, was administered one week later. The means and standard deviations on the retention test are reduced because subjects were asked to answer either the odd or even numbered items. Thus, a perfect score was 20, rather than 40.

Table 1

Means and Standard Deviations:
Post Test and Retention Test

		Post (40 questions)		Retention (20 questions)	
Groups	*N*	*M*	*SD*	*M*	*SD*
1	24	27.29	7.90	11.92	4.03
2	20	19.45	8.82	9.60	3.59
3	22	25.23	6.74	10.64	4.01
4	26	22.88	9.63	8.65	4.77
5	20	19.65	9.96	7.65	4.82
6	23	22.74	9.54	9.13	4.21

In order to get a clearer understanding of which posttest means were significantly different among the groups, the scores were analyzed via analysis of variance, followed by the Tukey HSD procedure. The *F* ratio was significant, *F* (5,129) = 2.81, $p < .03$, and one of 21 paired comparisons was significant. Group #1 (student-generated questioning from a summary, a pre-reading activity), was different from group #2 (student-generated statements following reading).

Retention means were similarly analyzed. The *F* ratio was significant, *F* (5,129) = 2.81, $p < .02$, and one of 21 paired comparisons was significant. Group #1 (student-generated questioning from a summary, a pre-reading activity), was different from group #5 (students read and study the story).

The significant post-hoc comparisons are in the direction predicted by the theoretical framework. However, these results need to be interpreted cautiously because of the likelihood of Type I error.

DISCUSSION

With four practice trials, all students were able to generate questions from a summary. Examination of the students' performance by levels indicated that only those students whose scores fell in the bottom fifth on the immediate and delayed recall had difficulty generating questions. The questions were not analyzed for this paper. However, the preliminary analysis of questions indicated that even poor readers improved over the four trials in their ability to generate questions.

This pilot study does suggest a facilitative effect on both immediate and delayed recall from student-generated questions. Where differences were found in immediate and delayed recall they were in favor of Group One. While the findings of this study are not conclusive, self-generated questioning does appear to be a viable technique worthy of further investigation with young readers.

REFERENCES

Allington, R. Improving content area reading instruction in the middle school. *Journal of Reading*, March, 1975.

Allington, R., Chodos, L. B., Domaracki, J., & Truex, S. Passage dependency for diagnostic oral reading tests. *Reading Teacher*, 1977, *30*, 369-75.

Ausubel, D. P. The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, 1970, *51*, 267-272.

Berlyne, D. E. An experimental study of human curiosity. *British Journal of Psychology*, 1954, *45*, 256-265.

Boker, J. R. Immediate and delayed retention effects of interspersing questions in written instructional passages. *Journal of Educational Psychology*, 1974, *66*, 96-98.

Duffy, R., Sherman, G., Allington, R., Roehler, L., & McElwee, M. *How to teach reading systematically*. New York: Harper, Row, 1973.

Frase, L. T. Boundary conditions for mathemagenic behaviors. *Review of Educational Research*, 1970, *40*, 337-347.

Frase, L. T. & Schwartz, B. Effect of question production and answering on prose recall. *Journal of Educational Psychology*, 1975, *67*, 628-635.

Levin, J. Comprehending what we read: An outsider looks in. *Journal of Reading Behavior*, 1971, *4*, 18-27.

Morse, J. M. *Effect of reader generated questions*. Paper presented at the National Reading Conference, 1975.

Natkin, G. & Stahler, E. The effects on adjunct questions on short and long term recall of prose materials. *American Educational Research Journal*, 1969, *6*, 524-536.

Pearson, D., & Nicholson, T. *Scripts, texts, and questions*. Paper presented at the National Reading Conference, Atlanta, 1976.

Peeck, J. Effect of prequestions on delayed retention of prose material. *Journal of Educational Psychology*, 1970, *61*, 241-246.

Rickards, John P. Processing effects of advance organizers in text. *Reading Research Quarterly*, 1976, *11*, 599-622.

Rothkopf, E. Z. Some theoretical and experimental approaches to problems in written instruction. In J. D. Krumboltz (Ed.), *Learning and the Educational Process*. Chicago: Rand-McNally, 1965, 193-223.

Rothkopf, E. Z. & Kaplan, R. Exploration of the effect of density and specificity of instructional objectives on learning from text. *Journal of Educational Psychology*, 1972, *63*, 295-302.

Schwartz, E. and Scheff, A. Student involvement in questioning for comprehension. *Reading Teacher*, November, 1975.

Smith, A. *The effectiveness of training students to generate their own questions prior to reading*. Unpublished doctoral dissertation, Syracuse University, 1972.

Smith, F. *Comprehension and learning*. New York: Holt, Rinehart and Winston, 1975.

Stauffer, R. *Directing reading maturity as a cognitive process*. New York: Harper, Row, 1969.

Tuinman, J. Determining the passage dependency of comprehension questions in five major tests. *Reading Research Quarterly*, 1973-1974, *9*, 206-223.

Tuinman, J. Inspection of passages as a function of passage dependency of the test items. *Journal of Reading Behavior,* Summer 1972-1973, *5,* 186-92.

Wiener, M. & Cromer, W. Reading and reading difficulty: A conceptual analysis. *Harvard Educational Review,* 1967, *37,* 620-643.

KRISTINA G. HAYWARD
VINCENT P. ORLANDO
EMERY P. BLIESMER
The Pennsylvania State University

EFFECTIVENESS OF A STUDY MANAGEMENT COURSE FOR "NON-TRADITIONAL" STUDENTS

The Developmental Year Program at The Pennsylvania State University included special English, Mathematics, Speech and Reading courses created to serve the needs of "non-traditional" students admitted under open admission policies. In 1975, Muia and Bliesmer (1975) reported an analysis of performance of these students during their first term and each of four successive terms. They found that the attrition rate of students steadily increased as they moved from remedial terms to successive non-remedial terms. Subsequently experience and observation of the present authors indicated that, although students improve and perform reading skills more adequately, apparently they fail to make use of these in regular course work. It seemed apparent that reading skill improvement alone was not sufficient for successful completion of university courses; functional change involving more efficient use of study habits also seemed essential.

Fox (1965) proposed a method for affecting more efficient use of study habits through application of behavioral principles. His study program was based upon the maximal use of available reinforcers, successive approximation, a fixed-ratio schedule of reinforcement, and a consistent set of stimuli. His preliminary attempts demonstrated sufficient success to warrant adaptation to the Developmental Year Program.

The purpose of the present study was to determine if the adapted program developed for Penn State's non-traditional students would: 1) affect their overall reading performance as measured by the Reading Power section of the *Iowa Silent Reading Test*, Level 11; 2) affect their overall academic performance as measured by Grade Point Average (GPA) change; and 3) affect their performance in the specific university courses used for applying the study management system.

METHOD

Subjects

Because performance in a regular university course was one of the variables being studied, registration for the four sections of the course was restricted to those Developmental Year students concurrently enrolled in a university course which involved use of a textbook. Therefore, students taking Mathematics or English Composition courses exclusively were not included. The fifty-four students, 30 males and 24 females, who registered for the course were informed at the outset of its experimental nature and of the course content, requirements, and evaluation procedures. Another group of 54 Developmental Year students not enrolled in the study management program, but enrolled in regular university courses, were randomly chosen as a control group.

In order to determine if both groups were essentially equivalent, their overall reading abilities were compared. A Behrens-Fisher t-test was applied with respect to the Reading Power section of the Iowa test. No significant difference was found between the two groups, $t(106) = .32, p > .05$. In an additional attempt to assess equality of the two groups, the students' predictive category numbers were also compared. The predictive number, assigned to each student upon admission to Penn State, was determined by using a formula containing SAT Verbal and Mathematics scores along with high school GPA. The results of a Behrens-Fisher t-test showed a significant difference, $t(87) = 2.38, p < .05$. This indicated that the members of the experimental group, with a significantly higher category number, would be less likely to succeed at Penn State. It was, therefore, concluded that the groups were equivalent on overall reading ability, but that the control was favored with respect to predicted success in college.

Materials and Procedures

The purpose of the study management program, then, was to provide instruction in, as well as opportunity for, consistent application of study skills, using regular university textbooks. To facilitate this purpose, modification of Fox's behavioral techniques to permit work in a group rather than an individual situation was completed. Students met three times a week for one hour in one of four sections of 10 to 15 students. Only the textbooks which the students were using in their regular university courses were used in the implementation of the SQ3R system. In order to maintain a consistent set of stimuli, time and location of the classes were fixed. Each student used his or her own textbook. In line with Fox's suggestions, short study sessions of approximately 10 to 15 minutes and breaks after completions of work segments were used as reinforcers. As rewards, students took two to three breaks during an hour session, initially; however, they soon reduced this to approximately one break per session, indicating willingness to complete larger work segments.

During the first two weeks, Robinson's (1962) SQ3R method was gradually introduced. Students began each new chapter of their textbooks by previewing and writing short summaries or outlines. On the basis of their summaries, students wrote questions which they thought would be answered in the text. They then read a segment of the chapter for approximately 15 minutes and subsequently wrote a summary of what they had read, with the book closed. In SQ3R terminology, the written summary would be equivalent to a recitation. As a final step students then opened their books and reviewed their summaries, adding any significant details or deleting irrelevant ones.

Analysis

In order to assess the effectiveness of the study management course on overall reading performance the Reading Power (Vocabulary and Comprehension) section of the Iowa Silent Reading Test, Level II, was administered before and after treatment. Differences in overall reading performance were measured through the application of a dependent *t*-test to pre/post Reading Power percentile scores.

The effect of the program on students' overall academic performance was measured by calculating the GPA (Grade Point Average) change of both groups from the preceding term to the current one. Because students received credit for the Study Management Program, the experimental group's post-GPA was computed with their reading grade excluded. An independent *t*-test was administered to compare the amount of change evidenced by both groups.

To measure the effect on students' performance in specific university courses, the mean grades of 31 of 54 experimental group students enrolled in introductory courses in Biological Science, Psychology or Sociology were compared with 31 control group students enrolled in the same courses. An independent *t*-test was performed on these data to measure the differential performance of the two groups.

RESULTS

A significant difference between pre and post test Reading Power scores indicated that overall reading gains were achieved by the experimental group, $t(53) = 3.66, p < .05$.

The mean GPA for both experimental and control groups declined, .09 for the experimental group and .36 for the control group. Although the experimental group's decline was smaller, the difference between groups was not significant, $t(106) = 1.51, p > .05$. The mean grade for the specific university courses studied was computed for both the experimental and control groups. The mean grade of 1.65 received by the experimental group was significantly higher than the control group's .87, $t(59) = 3.09, p < .05$.

DISCUSSION

The results indicated that the Study Management Program had a significant effect on students' performance in specific university courses and a positive but not significant effect on overall GPA. Significant overall reading test score changes were also indicated for the experimental group. However, conclusive support for the effectiveness of the program must remain tentative since post-test reading scores for (and consequently, comparison with) the control group were unavailable.

In attempting to determine the reasons for the effectiveness of the Study Management Program, two factors related to the study behavior of non-traditional students seemed to be most prevalent: a lack of sufficient study time and ineffectiveness of study behavior. Discussions with Developmental Year students indicated their resistence to long study hours. Studying only one or two hours for a mid-term examination, for example, was not uncommon. This in conjunction with their deficient background, often led to failure in university courses. Therefore, the fact that students in the experimental group were gradually introduced to consistent hours of study time would seem to have had a facilitative effect.

Secondly, although most students would agree that effective study results in a better understanding of the material, few actually measure the effectiveness of their studying through some type of post-recitation. The students participating in the course, however, regularly tested themselves by closing their books and summarizing the new material they had read. This systematic closed-book recitation provided not only additional rehearsal time for new material but also supplied the students with an opportunity to organize this new material into meaningful summaries.

In summary, the program appeared to be successful in affecting change in students' study behavior, with success stemming from a regular pattern of studying, use of an effective study system which involved recitation, and the use of relevant academic materials for applying these principles.

REFERENCES

Fox, L. Effecting the use of efficient study habits. In R. Ulrich, T. Strachnick, and J. Mabry (Eds.), *Control of human behavior.* Glenview, Illinois: Scott Foreman Co., 1965.

Muia, J. A. & Bliesmer, E. P. Prediction of academic success of special groups of university freshmen. In George H. McNinch and Wallace D. Miller (Eds.), *Reflections and investigations on reading.* Twenty-fifth Yearbook of the National Reading Conference, 1976, 88-96.

Robinson, F. P. *Effective reading.* New York: Harper and Row, 1962.

31

MARILYN M. FAIRBANKS
CAROLYN CASTELLO
West Virginia University

MEASURING AND EVALUATING CHANGES IN UNDERLINING AND NOTE-TAKING SKILLS

Note-taking and underlining skills have long been considered important aspects of college reading-study skills programs designed to help students improve in academic efficiency and performance. As part of an intensive study of sixty-nine reported college programs using academic performance as an evaluation criteria, Fairbanks (1973) tabulated the reported inclusion of twenty-eight different skills. "Note-taking and/or underlining" were reported as included in thirty-two of the sixty-nine programs, a higher incidence of reported inclusion than that for any other skill. The high incidence of inclusion of note-taking and/or underlining procedures in current materials designed to aid college students in study improvement is another indication that these skills are considered important. Examples include materials by Baldridge (1972), Gilbert (1955), McGraw-Hill Basic Skills System (1970), Morgan (1969), Mortensen (1975), Pauk (1974), Robinson (1962), Smith (1966), Smith (1970), and Staton (1966).

Despite the high level of acceptance given note-taking and underlining techniques as important components of college reading-study skills programs, systematic research regarding improvement in note-taking and underlining skills, or the influence of such skills on academic performance, has been extremely limited. Palmatier (1971) reported that "only five experimental efforts related to note-taking techniques have been reported since the turn of the century." No experimental studies were found dealing specifically with underlining techniques. Both Paulsen (1968) and Cooper (1969) deplored the state of research in the study skills area; Paulsen cited the lack of appropriate measuring instruments as one of the difficulties.

In those studies that were located, improvement in some comprehension measure was used as a criterion in a number of studies. For example, a study involving comparison of the effects of simultaneous versus subsequent note-taking from lecture on immediate and delayed learning (one week later), Van Matre et al (1975) used comprehension scores on prepared questions as an indicator. Also using prepared comprehension questions as the criterion, Van Matre and Carter (1975) found that taking notes without review resulted in poorer test results than when notes were not taken at all. Accuracy and completeness of free recall statements to compare the effectiveness of note-taking methods were used by DiVesta (1972) and Schulz (1972).

The one study that involved a measuring scale for evaluating the quality of the notes directly was that of Palmatier (1971) in which ratios for completeness, clarity, and conciseness of notes had been devised.

In a cooperative pilot "dropout intervention" project at WVU involving the University Reading Laboratory staff and facilities, and Psychology Department personnel, note-taking and underlining were included in a twelve-session study skills program (Feindler, 1977). Procedures were worked out for the teaching of these skills, and the authors of this paper devised scales to measure performance and improvement in note-taking and underlining skills. Forty-nine second semester freshmen were involved in the study skills sections, and twenty-four second semester freshmen served as a non-treatment control group.

PROCEDURES AND RATING SCALES

Note-taking

A taped lecture was used for note-taking. The note-taking scales devised were designed to evaluate progress on the note-taking method advocated, which involved the taking of organized raw notes during the lecture, and the subsequent refining of notes including marginal clues. Performance and improvement in note-taking were measured by both the quality of notes produced and responses to prepared multiple-choice questions. The scale devised, and point values indicated, is included as Table I.

Table I
Note-Taking
(10-point scale)

	Criteria	0	1	2
A.	Raw notes; differentiation between main points and details by indentation, spacing, and/or marking systems	None	Scant	Definite pattern of differentiation
B.	Relationship of subtopics to headings	No subtopics or not related	Inconsistent	Subtopics consistently related to headings
C.	Use of cue words (In addition to headings)	No cue words	Too many or too few	Adequate use of cue words
D.	Separation of cue words from notes, and balance of cue words to notes	No separation of cue words from notes (or no cue words?	Indefinite cue words and/or inadequate balance	Adequate separation and balance
E.	Phrasing (meaning conveyed; not over-wordy)	Meaning not clear and/or over-wordy	Inconsistency in conveying meaning and/or tendency toward excessive verbage	Meaning conveyed with economy of words

Underlining

The underlining method used advocated a modest amount of underlining which included both main ideas and major details, and was generally similar to that of Wark (1970). Underlining performance was judged solely by the scale described generally in Table II. No comprehension check was used.

Table II

Underlining
(6-point scale)

	Criteria	0	1	2
A.	**AMOUNT**	**Like models A or C*** **(less than 25% or more than 75% underlined, or whole sentences underlined)**	**Between two models**	**Like Model B***
B.	**COMPLETENESS**	**Half or less of major points covered**	**3/4 or more of major points covered**	**All major points identified covered**
C.	**ACCURACY**	**Underlining conveys erroneous information**	**Underlining confusing**	**Underlining accurately coree major ideas**

*For models, students were referred to: Wark and Mogen, 1970, p. 14.

METHOD AND RELIABILITY OF JUDGING

All student identification was removed from papers to be rated. One selection was used for note-taking both as the pre and post performance indicator; another selection was used for underlining for both pre and post performance. Before papers were read for rating, they were coded and pre and post papers as well as experimental and control group papers were intermixed as one group. Two judges rated the note-taking and underlining samples. General criteria described in Tables I and II were further formalized after the preliminary checking of five sample papers for both scales by both judges. These papers were mixed with the others for subsequent rating. During the rating process itself, each judge read four papers and the ninth was read by both judges. During this process, formalization of criteria was reviewed and further defined as necessary. When a difference in rating between judges existed, the rating used was an average of the two.

With reference to rating of student underlining, on the twenty trial papers rated by both judges, identical results were obtained in fifteen (75%) of the papers; judges varied one point in five (25%) of the cases.

With reference to rating of student note-taking from lecture, complete agreement of judges was achieved in sixteen (80%) of the twenty cases. Judges differed by one point in three instances(15%) and by two points in one instance (5%).

RESULTS

Note-taking scores on the 10 point note-taking scale described in Table I ranged from 0 to 7 (pre) and from 3 to 10 (post) for the study skills groups, and from 2 to 8 (pre) and 3 to 6 (post) for the control groups. Mean note-taking scores for the study skills groups increased from 4.75 to 6.45 on pre-post measures, with standard deviations of 1.51 and 2.07 respectively. The "incidental" control group decreased slightly on pre-post note-taking means, from 5.71 to 5.57, with pre-post standard deviations of 1.09 and .91. Between-group differences were not significant. An analysis of variance on within-group change in note-taking measures for the experimental groups indicated that the pre-post difference was significant at the .001 level.

Comprehension of the material on which notes had been taken was evaluated by a 50-point multiple choice test. Scores for the study skills students ranged from 10 to 40 (pre) and from 20 to 45 (post). Control group scores ranged from 10 to 40 (pre) and 10 to 45 (post). Pre and post-test comprehension means for the study skills groups were 25.65 and 31.49 respectively, with standard deviations of 7.72 and 6.67. Control group means were 26.04 (pre) and 27.14 (post) with standard deviations of 9.47 and 7.52. Between-group differences were not significant. Analysis of variance of within-group change in comprehension scores for the study skills groups indicated a difference significant at the .001 level.

With reference to changes on the 6 point underlining scale included as Table II, scores ranged from 0 to 6 for both experimental and control groups on both pre and post measures. Mean scores for the study skills groups were 3.13 and 3.68 with standard deviations of 1.[illegible]1 and 1.53 respectively. Control group pre-post scores showed a slight decrease in means, from 3.08 to 2.98, with standard deviations of 1.39 and 1.36. Neither between-group or within-group differences reached significance at the .05 level or beyond. The probability of significant difference relative to within-group change for the study skills groups was .07.

IMPLICATIONS

The use of rating scales, such as those described, by judges who further refined and formalized rating scale criteria to standardize judgments, could be used in a number of ways. Comparisons between the effectiveness of note-taking and underlining could be made; the relationship between note-taking and underlining with academic achievement and performance could be studied. Investigators might well find it helpful to revise scales, or to design completely new ones, as appropriate for their own particular objectives and teaching methodologies.

However, of even greater interest is the improvement of comprehension as a function of study techniques. The value of more sensitive scales for evaluating note-taking and underlining behavior may well be realized in future studies which examine, both experimentally and correlationally, the relationship between these commonly used study techniques and comprehension.

REFERENCES

Baldridge Reading Instruction Materials, Inc. *Reading and study techniques for academic subjects.* Greenwich, Conn.: Baldridge Reading Company, 1972.

Cooper, B. & Fox, J. M. Students' study habits, attitudes, and academic attainment. *Universities Quarterly,* 1969, *2,* 203-212.

DiVesta, F. J. & Gray, G. S. Listening and note taking. *Journal of Educational Psychology,* 1972, *63,* 8-14.

Fairbanks, M. M. *An analytical study of the relationship of specified features of reported college reading improvement programs to program effect on academic achievement* (Doctoral dissertation, West Virginia University, 1972) Ann Arbor, Mich.: University Microfilms, 1973, No. 73-12, 938.

Feindler, E. *A comparison of two service delivery models in the study skills training of potential college dropouts.* Unpublished masters thesis, West Virginia University, 1977.

McGraw-Hill Basic Skills System. *Listening and notetaking.* Heightstown: McGraw-Hill, 1969.

Morgan & Deese, *How to study,* Heightstown: McGraw-Hill, 1969.

Mortensen, D. R. *The lazy student's way to excel.* Mortensen Pub., 1975.

Palmatier, R. A. Comparison of four note-taking procedures. *Journal of Reading,* 1971, *14,* 235-258.

Pauk, W. *How to study in college* (2nd ed.). Atlanta: Houghton-Mifflin Company, 1974.

Paulsen, S. C. *The scientific basis of our knowledge about study methods.* Copenhagen: The Danish Institute for Educational Research, 1968.

Schulz, C. B. & DiVesta, F. J. Effects of passage organization and note taking on the selection of clustering strategies and on recall of textual materials. *Journal of Educational Psychology,* 1973, *72,* 244-252.

Smith, D. F. *Learning to learn.* Chicago: Harcourt, Brace and World, Inc., 1966.

Smith, S. *Best methods of study.* New York: Barnes and Noble, Inc., 1970.

Staton, T. *R.S.V.P.: A dynamic approach to study.* Glenville, Ill.: Soctt, Foresman, and Co., 1966.

Van Matre, H. H. et al. *Learning from lecture: Investigations of study strategies involving note taking.* (ERIC Document Reproduction Service No. ED 120 684)

Van Matre, N. H. & Carter, J. F. *The effects of note-taking and review on retention of information presented by lecture.* (ERIC Document Reproduction Service No. ED 120 684)

Wark & Mogen. *Read, underline, review.* Heightstown: McGraw-Hill, 1970.

STEVEN M. BIALICK
DIANE H. BIALICK
DAVID M. WARK
University of Minnesota

THE EFFECTS OF EXPERIMENTER-UNDERLINING AND EXPERTISE ON THE RETENTION OF TEXTUAL MATERIAL

On college campuses throughout the United States, used textbooks are circulated containing the underlinings of the previous owner. Fowler and Barker's survey (1974) of 200 randomly sampled used texts found 92% had "significant applications" of emphasizing including underlining and highlights. What are the effects of these markings? Are they helpful or distracting? The anecdotal folklore is inconsistent. Some students actively seek out books that have been previously underlined, as if they would be helped. Other students look for clean copy, preferring to do it themselves. But what in fact is the empirical evidence regarding previously underlined textbooks? Are they helpful or not?

There is experimental literature on the effects of previously underlined texts. In his review Browning (1976) summarizes 12 control group studies of underlining. Six were significant and positive, two showed non-significant trends, and four were clearly negative. One of the early tests of experimentor — underlined material (Christenson and Stordahl, 1955) found no significant increase in comprehension for the underlined group. This lack of effect was replicated by Hershberger (1964), Rocchio (1972), and Earp (1959). On the other hand, subjects reading experimentor-underlined materials did score significantly higher on test items dealing with the underlined material in studies by Cashen and Leicht (1970), Leicht and Cashen (1972), and Schnell and Rocchio (1974).

The mixed results lead to questions of why underlining works, and what variables produce better learning from previously underlined materials.

The effects of underlining seem to be an example of a phenomena identified by von Restorff (1933). He found that isolating a single word against a homogeneous background of other words facilitated recall of a target word. Wallace (1965) reviewed the studies of the von Restorff phenomena. He concluded that the isolation effect is a reliable one, and better recall of the isolate occurs at the expense of the remaining list units. Underlining, whether with a simple lead pencil or dramatic color printing, would have the effect of isolating certain words, phrases, or sentences and make them stand out from the page. However, in light of the inconsistent findings reported above, there must be variables which can either weaken or enhance the von Restorff effect. One such variable may be the credibility of the source that produced the underlining.

After reviewing the literature on source credibility, Insko and Schopler (1972) concluded that a highly creditable source is perceived as both expert and trustworthy. Such a source possesses greater persuasive impact. One might expect the von Restroff effect to be magnified when the underlining is perceived as being that of an expert, and diminished when perceived as having been done by less competent readers. In the present study the credibility of the source of the underlining was experimentally manipulated. Underlining was presented as either the work of low expertise (incoming University freshmen) or high expertise (law students).

The literature review provides the basis for 3 predictions:

1. Underlined material that is relevant to a test will facilitate learning and lead to a higher measure of retention.
2. Underlined material that is irrelevant to a test will be learned at the expense of relevant material and lead to reduced retention.
3. Both facilitating and reducing effects of underlining will be enhanced by increasing the perceived prestige of the source of the emphasis.

METHOD

Subjects

Subjects were 81 high school seniors from four psychology classes at a suburban Minnesota high school Nearly all of the students were considered by their counselor to be college bound.

Procedures

Intact classes were randomly assigned to a high or low prestige condition. The students were tested during their regular class hours, in their psychology classroom. They were told that they were taking part in an experiment. The students read a mimeographed copy of chapter three from *A Short History of 16th Century England* (Woodward, 1963). The text has a Fry readability of eleventh grade (Fry, 1968). Thirty-two multiple choice four alternatives test items were available for chapter three. This test had been validated in previous research (Wilmore, 1966). For this experiment, a sub-test of sixteen items was composed of questions that could be answered from material underlined in one treatment condition.

Within intact classes students were randomly assigned to one of three types of reading material. One treatment (relevant underlining) contained sixteen evenly distributed penciled underlinings with sufficient information highlighted to correctly answer 16 criterion items; a second treatment (irrelevant underlining) contained distributed underlinings which isolated information completely irrelevant to any of the questions on the test; the third treatment (no underlining) contained no marking whatsoever. Underlinings for the two experimental conditions were of equal length and style, with a density of approximately one per page. The presence of underlining in some reading copies was explained by the "aside" in the following instructions to the classes:

> (aside): Oh, and please don't make any marks in the reading or test booklet. The test was just given to (University freshmen/law students) and some marks were made. This material has to be used again, so please, don't make any marks in it.

> Now turn over the reading booklet and begin reading. You have 30 minutes.

Two classes were told university freshmen had just taken the test (Low Prestige), and two were told law students had just taken it (High Prestige). This represented the credibility manipulation as to the perceived expertise of the underliner. A check question was written on the board to determine if the manipulation worked:

> Before coming to your high school, this test was given to:
> 1) Law students
> 2) University freshmen
> 3) Housewives
> 4) No one

All but two of the students correctly answered this question.[1]

RESULTS

Table 1 reports the means and standard deviations for each treatment condition. Although 81 students were in the initial sample, not all of them appear in the final report. Three were eliminated for improperly filling out their answer sheet. Six were randomly removed in order to equalize the number in each prestige X condition cell. The final *N* equals 72.

Table 1

Means and standard deviations of six groups' retention following various levels of underlining and prestige. *N* = 12 per cell

	Type of Underlining					
	Relevant		None		Irrelevant	
	M	*SD*	*M*	*SD*	*M*	*SD*
Hi Prestige	8.91	2.90	5.75	1.28	4.00	1.41
Low Prestige	6.58	2.64	5.66	1.61	4.66	1.49

Table 2 reports the two way Anova. Differences between the types of underlining were significant at well beyond the .01 level. Levels of prestige of the underliner were not significant. However, there was a significant interactions score, suggesting that some combinations of main effects were more effective than others.

---

[1] Copies of the instructions, test items, keys, and a master copy of the relevant and irrelevant underlining are available for the 3rd author at University of Minnesota, 1420 Eckles Ave., St. Paul, Minnesota 55108.

Table 2

ANOVA of six groups' retention following various
levels of underlining and prestige. $N = 72$.

Source	Sum of Squares	*df*	Mean Square	*F*
Underlining	141.86	2	70.93	17.77**
Prestige	6.125	1	6.125	1.53
Interaction	29.25	2	14.62	3.66*
Error	263.41	66	3.99	

*$p < .05$
**$p < .01$

Specific predictions were tested using the Scheffe technique (1953) to evaluate the differences found after the Anova. In essence, the technique adjusts probability levels to provide correct statements of the chance of particular findings.

Prediction 1 was that relevant underlined material would be retained better than unmarked text. The prediction was tested by evaluating the difference between mean scores on the 16 underlined items for the Relevant Underlining and No Underlining treatments. For the High Prestige group, the mean for Relevant Underlining ($M = 8.91$) was greater than for the No Underlining condition ($M = 5.75$) at the $p < .01$ level. The first prediction of facilitating effect for Relevant Underlining is confirmed for those students who thought they were working with material underlined by law students, presumably high prestige source. And the direction of these results, while not significant, is as expected for those students reading papers that they thought were underlined by college freshmen.

Prediction 2 was that Irrelevant Underlined material would lead to lower retention than No Underlining. According to the Scheffe test, none of the differences between No Underlining and Irrelevant Underlining were significant at $p <$ 05. However, as for prediction 1, the trend of the data is in the desired direction; the means for Irrelevant Underlining are both lower than for No Underlining.

There may have been a floor effect that mitigated against finding a greater difference. Students in the Irrelevant condition earned scores at chance level (¼ of 16 items). It would be unlikely for them to earn still lower scores, and show a greater effect. (The authors wish to acknowledge Dr. Richard C. Anderson's kindness in pointing out the salutary implication of the possible floor effect.)

The third prediction involved the differential effect of high and low prestige sources. According to the Anova results, there were no effects attributable directly to level of prestige. However, there is a significant interaction, which appears to be between High Prestige and Relevant Underlining. Thus in the Relevant condition, the High Prestige source seems to be more effective than the Low Prestige. The Scheffe test confirms the locus of the significant interaction. The retention for Relevant Underlining and High Prestige ($M = 8.91$) is significantly greater than for Relevant Underlining in Low Prestige ($M = 6.6$) at $p < .05$ level.

And again, there does seem to be trendish evidence for a prestige effect. While the High Prestige Relevant Underlining is higher than that for Low Prestige, the High Prestige Irrelevant Underlining is lower than for Low Prestige. When the means for the 6 cells are plotted, the trend lines are not parallel. They cross, supporting the notion that High Prestige may have an exaggerating effect on underlining. However, the directions are not statistically significant.

DISCUSSION

The present study adds something to the literature of underlining. First, the von Restroff effect appears to be upheld. Material that was well underlined, merely by a lead pencil, was well learned. In the Relevant Underlining group, learning showed up as higher scores on the 16 highlighted items. But the von Restroff phenomena would predict another effect, which also showed up. When irrelevant material is underlined, it is retained at the expense of the other material in the text. Consequently, the students in the Irrelevant Underlining group tended to retain less than the students who read material without any underlining at all. Von Restroff phenomena seems to handle the general direction of the data.

The other variable investigated in this paper was the prestige of the presumed underliner. The impact of High Prestige is to magnify the effects of Relevant Underlining. Perhaps the negative results reported in the underlining literature were due in part to levels of prestige. Christenson and Stordahl, for example, had air force trainees reading material that was structured in a variety of ways. They used outlines, summary statements, embedded questions, and underlining as a technique for aiding recall. These cues were all printed as part of the regular material. Hence they all had approximately the same prestige effect. And there were no observable differences between them. Unfortunately, the authors did not use a control condition by presenting straight text material. What was presented as negative findings may in fact be non-differentiated positive findings.

In the studies where underlining was effective it is also possible to find evidence for a prestige effect. In the Leicht and Cashen studies the reading materials were printed magazine articles with certain words underlined in red. These underlinings, clearly not part of the regular material, were implicitly endorsed by the highest possible prestige source, the instructor. And in those conditions, they were effective.

One further point should be raised about the prestige effect. In some studies where the learning was effective, the underlined materials were seemingly added to the original documents. For Christenson and Stordahl, for example, the typographical cues were built in by the printer. The same was the case for Hershberger. In the Leicht and Cashen and Cashen and Leicht studies the underlinings were apparently add-ons provided by the instructor. Thus the effect of the emphasis was augmented by source prestige. It may be that part of the effectiveness of underlining is due to it being seen as an add-on, something special, not included in the printing. Preprinted underlining may be something like an appendix or preface: presumably useful but generally ignored by students. If so, expensive multi-color printing and typographical gimmicks in textbooks may be less effective than publishers believe. It may be that part of an underlinings value comes from its perception of being something extra added on and special.

REFERENCES

Bialick, D. *Use of the twenty statements test to predict individual differences in reliance on underlining.* Unpublished summa thesis, University of Minnesota, Minneapolis, 1975.

Browning, W. J. A critical review of research and expert opinion on the underlining study aid. In G. Miller & E. McNich (Eds.), *Reflections and investigations on reading.* 25th National Reading Conference Yearbook, Clemson, S.C.: National Reading Conference, 1976.

Cashen, V. & Leicht, K. Role of the isolation effect in a formal educational setting: Underlined and non-underlined printed material. *Journal of Educational Psychology,* 1970, *61,* 484-486.

Christenson, C. & Stordahl, K. The effect of organizational aids on comprehension and retention. *Journal of Educational Psychology,* 1955, *46,* 65-74.

Earp, N. *The effect of certain study conditions on the achievement of college freshmen.* Unpublished Ed.D. dissertation, Colorado State College, Greeley, 1959.

Fowler, R. L. & Barker, A. S. Effectiveness of high lighting for retention of text material. *Journal of Applied Psychology,* 1974, *59,* 358-364.

Hershberger, W. Self-evaluational reading and typographical cueing: Techniques for programming self-instructional reading materials. *Journal of Educational Psychology,* 1964, *55,* 288-296.

Insko, C. A. & Schopler, J. *Experimental social psychology.* New York: Academic Press, 1972.

Leicht, K. & Cashen, V. Type of highlighted material and examination of performance. *Journal of Educational Research,* 1972, *65,* 315-316.

Rocchio, D. *The effect of prior underlining on immediate comprehension of reading passages.* Unpublished independent study project, University of Missouri, St. Louis, 1972.

Scheffe, H. A method for judging all contrasts in the analysis of variance. *Biometrika,* 1958, *40,* 87.

Schnell, T. & Rocchio, D. A comparison of underlining strategies for improving reading comprehension and retention. In E. McNich and G. Miller (Eds.), *Reading: Convention and inquiry.* 24th National Reading Conference Yearbook, Clemson, S.C.: National Reading Conference, 1976.

von Restorff, H. Uber die Virkung von Bereichsbildungen im Spurenfled. *Psychologie Forschung,* 1933, *18,* 299-342.

Wallace, W. P. Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin,* 1965, *63,* 410-424.

Willmore, D. *A comparison of four methods of studying a college textbook.* Unpublished doctoral dissertation, University of Minnesota, Minneapolis, 1966.

Woodward, G. W. O. *A short history of 16th century England: 1485-1603.* New York: Mentor, 1963.

JEROME C. HARSTE
CAROLYN L. BURKE
Indiana University

A NEW HYPOTHESIS FOR READING TEACHER RESEARCH: BOTH *TEACHING* AND *LEARNING* OF READING ARE THEORETICALLY BASED

Because of our involvement in the teaching and supervision of college students within reading practicum experiences, we have constant entry to a number of public school classrooms. One exciting dimension of this experience is the exploration of the teacher's role in assisting children with their acquisition of reading competency. What has become both readily apparent and surprisingly persistent concerning the relationship between reading instruction and the reading process is that: (1) despite atheoretical statements, teachers are theoretical in their instructional approach to reading, and (2) despite lack of knowledge about reading theory, per se, students are theoretical in the way in which they approach learning to read. Both students and teachers exhibit behavior which is sufficiently systematic to allow inferences about a theory which must underly that behavior.

To support this finding, observations made of both pupil and teacher behaviors are presented. The paper includes a research design for the continued exploration of this hypothesis and is intended to refocus inquiry in the area of reading teacher education research.

Before defining by example what is meant by theoretical orientation as observable in student and teacher behavior, a verbal definition of the concept seems in order. Put simply, a theory is a system of assumptions through which experiences are organized and acted upon. In terms of cognitive psychology (see Anderson, *et al.*, 1976), a theoretical orientation is best thought of as a cognitive structure or generalized schemata which governs behavior. Operationally, a theoretical orientation is a particular knowledge and belief system held toward reading. In practice, this knowledge and belief system operates to establish expectancies and strongly influences a whole host of decisions made by teachers and pupils relative to reading.

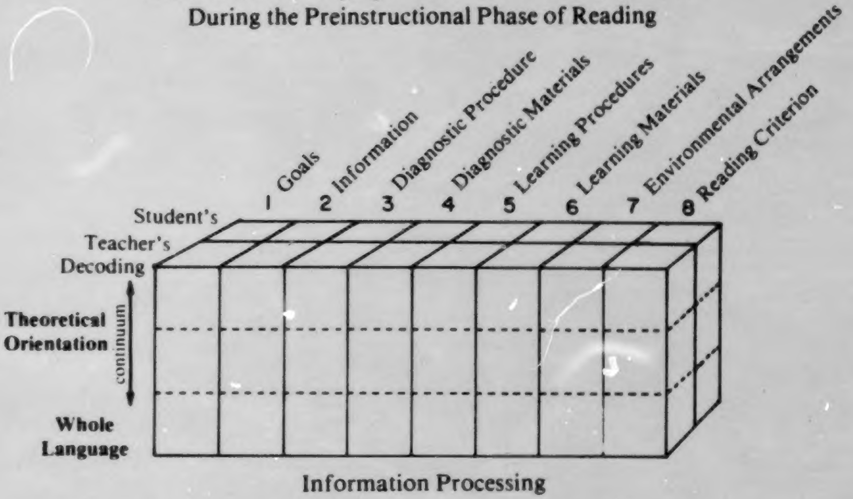
With this definition in mind, it might be well to restate the findings which we are reporting here; namely, that we found *both teachers and students* to have distinctive and identifiable theoretical orientations to reading and that once we had identified these orientations, subsequent reading performance and classroom behavior was found consistent with the model from which the person was operating.

The research paradigm used to explore these notions is presented in Figure 1. This paradigm suggests that a study of the decisions made by both teachers and pupils relative to what goals they select, what information or data they use to make decisions as to progress in reading, and the like, are key to the identification of theoretical orientation. It should be noted that this paradigm includes not only a teacher dimension, but also a student dimension. This is so because our field findings strongly suggest that student performance is often key to understanding teacher performance. Put simply, we have found student reading performance, at least in part, to mirror instruction. Student performance, in this sense, is an artifact of instructional history. An instructional setting allows the student to either reinforce or create an alternate schema. Put another way, students' predisposition to apply one theoretical model over another will be influenced by the instructional environment, i.e., the teacher's theoretical orientation or model of reading, or the text author's theoretical orientation or model of reading.

Before leaving this paradigm, it should be noted that what our field research has shown is that students and teachers operate out of one orientation or another consistently across information processing or decision points. If one were to build a profile for any given student or teacher from the data we have collected, one would end up by filling in all of the cells horizontally, i.e., indicating a consistent orientation across information processing or decision areas.

FIGURE 1.

A Schema for Studying Teacher and Student Behavior
During the Preinstructional Phase of Reading

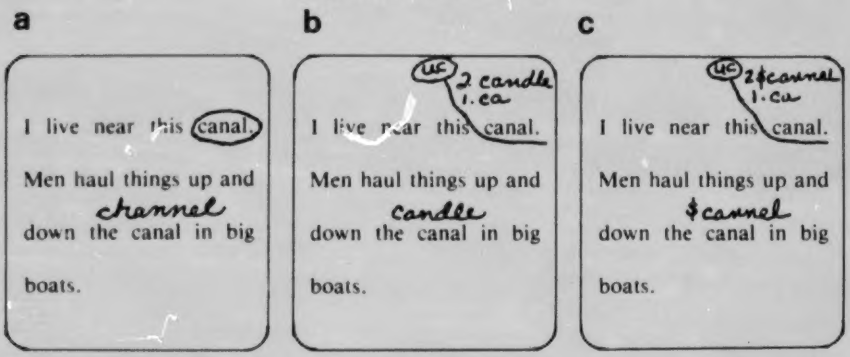

1 Goals
2 Information
3 Diagnostic Procedure
4 Diagnostic Materials
5 Learning Procedures
6 Learning Materials
7 Environmental Arrangements
8 Reading Criterion
Student's
Teacher's
Decoding
Theoretical Orientation
continuum
Whole Language
Information Processing

1. Goals Selected and Weighing of Goals
2. Information Selected for Diagnosis and the Weighing of such Information
3. Diagnostic Procedures To Be Used
4. Diagnostic Materials To Be Used
5. Learning Procedures To Be Used
6. Learning Materials To Be Used
7. Environmental Arrangements To Be Used
8. Reading Criterion To Be Used

THE INSTRUMENTS (SAMPLES)

To explore this hypothesis several instruments have been developed. Many of these instruments have been designed to simultaneously explore one or more information-processing or decision areas. For example, pupil information cards have been designed to represent pupil behavior reflecting a decoding, a skills, and a whole language orientation to reading. An example of one set of cards is presented in Figure 2. The marking system is explained to the teacher during an interview along with an actual reading of the card by the examiner. Teachers are asked to choose who they perceive to be the best reader. Because of the way in which the experiment is designed, the teacher's act of selection permits easy determination of theoretical orientation. A follow-up interview with the teacher is taped for the purpose of validating our classification of pupil information cards.

FIGURE 2
Pupil Information Cards



Code: canal = omission

channel
canal = substitution

uc = uncorrected miscue

2. cannel
1. ca = first attempt; second attempt

$ = sound like

a. oral reading sample - whole language orientation
b. oral reading sample - skills orientation
c. oral reading sample - decoding orientation

Similarly, a student interview form has been developed which attempts to identify a student's orientation toward reading through a variety of open-ended questions such as:

1. When you are reading and you come to something you don't know, what

do you do?

2. Who's the best reader you know? What does he (or she) do that makes him (or her) such a good reader?
3. If you were going to help someone with his reading, how would you help him?

It should be noted that the unit of focus (letters, words, or larger than word units) is left completely up to the subject to select. This selection, in and of itself, is key to understanding orientation to reading.

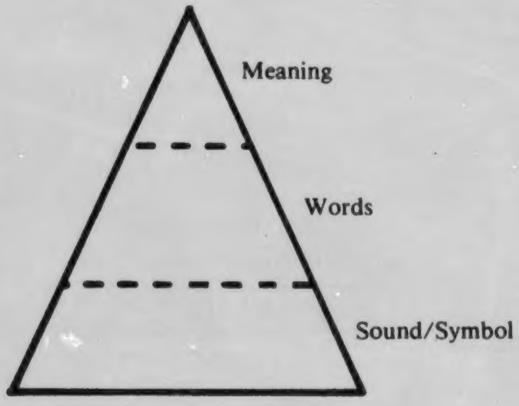
While we have developed many other research instruments and techniques for the exploration of this hypothesis, further elaboration will have to await some longer publication. These examples will serve to give the reader a feel for the type of procedures being used and the kinds of data being collected.

MAPPING DATA AGAINST POSITED THEORETICAL POSITIONS

To demonstrate the practical differences that theoretical orientation makes to both instruction and learning, the remainder of this paper will identify three current reviews of reading, specify the relationships between components in the reading process under each orientation, quote text book authors whose material reflects these representative views of reading, and finally conclude by presenting pupil and teacher data which demonstrate the practical implications of various models of reading in operation.

Sound/Symbol or Decoding Orientation. In this instance reading is perceived as an offshoot of oral language, the chief accomplishment of which is dependent upon developing and manipulating the relationships between the sounds of speech and their graphic symbols. While people who hold this view of reading don't argue against the existence of syntax and meaning as components of language, they do not see them as primary factors in the acquisition of the process. Language is perceived as a pyramid, the base of which is sound/symbol relationships, the capstone of which is meaning. Figure 3 illustrates this model.

FIGURE 3
Decoding Model of Reading



McCracken and Walcutt in the Teacher's Edition of *Basic Reading* (Lippincott, 1963), exemplify this orientation:

> "Do you purpose to define reading as mere word-calling without regard for meaning?"
>
> "Yes we do. Reading is, first of all, and essentially, the mechanical skill of decoding, of turning the printed symbols into the sounds which are language."

Teachers who exemplify this position chose Card C as the best reader (see Figure 2) because of the high grapheme-phoneme correspondence evident between the expected and the observed response. Readers operating out of this model demonstrated a concern with grapheme-phoneme correspondence production in their reading. The following coding worksheet is an example of a student operating out of a decoding model when reading:

\$ Cliver \$Turntell
The Clever Turtle

\$villigo
One day a man left his village to tend to his field of corn. But he

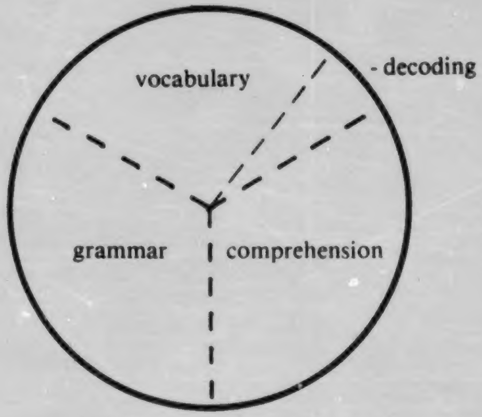
strite
found only an open place where young corn had stood ^ straight and

tall.[1]

KEY
Cliver
Clever = substitution
\$ = sounds like

Skills Orientation. A second orientation views reading as one of the four language arts — listening, speaking, reading, and writing. The four language arts are seen as being composed of (and thus learned as) a collection of discrete skills while sharing "common abilities."

FIGURE 4
Skills Model of Reading



Because language is perceived as a pie from which individual "skill slices" can be extracted for instruction, it becomes a relevant task to develop skill hierarchies. Text book authors who operate out of this model usually provide instruction in all three component areas for each lesson. Because persons holding this model believe the distinctive feature or key to reading success is the word, new vocabulary items are typically introduced prior to reading. Following silent and oral reading a series of comprehension questions are given. Workbook activities complete the model by providing skill practice on usage. Robinson, Monroe, and Artley in the *New Basic Readers* (Scott Foresman, 1962), best illustrate this model and conclude in the *Teachers' Edition* that one make sense out of reading by stringing words together. The quote which follows captures this orientation's emphasis on words as well as the notion that reading is a sequential skill mastery process.

> "Initially a child must learn to identify printed individual words and relate them to a meaningful context. This is best done by a) rapidly developing a basic sight vocabulary and b) teaching word recognition skills."

Teachers operating out of this model of reading select Card B as the best reader (see Figure 2) because of the child's production of a whole meaning word with good phoneme-grapheme correspondence.

Readers operating out of the skills model demonstrate a whole word miscuing pattern as the following example so vividly illustrates:

you star started
But he found only an open place where young corn had stood straight
tell
and tall.[2]

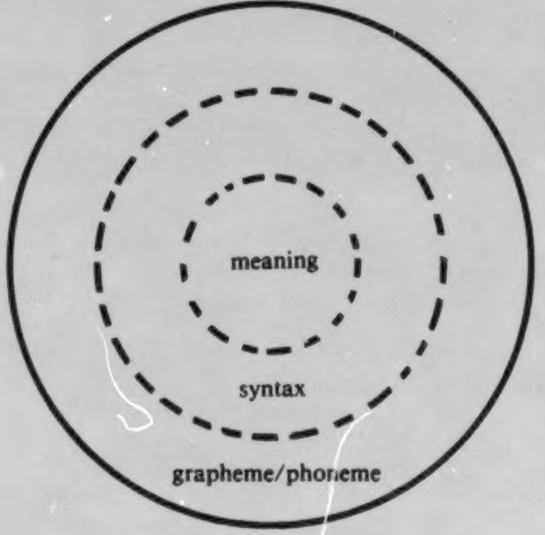
KEY
you
young = substitution

Please note that this reader, Jimmy, substituted only whole words while reading. While not all children operate this consistently, we have found one distinctive feature of a skills reader is his regular production of whole words as opposed to graphophonically similar nonsense words. Although we have not totally explored this notion, we have some initial data which suggests that the words substituted come largely from sight word lists introduced to the child. If this trend is found throughout all of the data we have collected, we will have clear evidence of the effect of instructional history upon student theoretical orientation.

Whole Language Orientation. A third theoretical orientation, which we term a whole language or language based approach, views reading as one of four ways in which the abstract concept of language is realized. This orientation assumes not only that the systems of language are shared, but that they are interdependent and interactive aspects of a process. Figure 5 illustrates that under this model language is conceived of as a sphere. This sphere is composed of a meaning core enwrapped in a syntactic structure and sheathed with a phoneme-grapheme system. When aspects of language are focused upon for instructional purposes,

FIGURE 5

Whole-Language or Language Based Model of Reading



the sphere is penetrated and all three systems are extracted simultaneously. In this view, reading, whether or not for instructional purposes, is always focused upon comprehending. Text book authors who compose materials from this perspective often do so building from the oral language base of the reader. Under this view speaking differs from reading only by the addition of the grapheme component in the outer ring of the model. Given this perspective, it follows that reading educators ought to build upon the strong language systems which the child already has mastered when teaching reading. *Scott Foresman Reading Systems* (Aaron, *et al.*, 1971), the *Sounds of Language Program* (Martin and Brogan, 1972), and the core of the *LEIR Program* (Van Allen, 1974) exemplify this approach. Kenneth Goodman, one of the authors of the *Scott Foresman Reading Systems Program,* clearly exemplifies this model:

> "Reading is the active process of reconstructing meaning from language represented by graphic symbols (letters), just as listening is the active process of reconstructing meaning from the sound symbols (phonemes) of oral language." (Smith, Goodman, and Merdith, 1976)

Teachers operating out of this model chose Card A as the best reader (see Figure 2), indicating that the child was willing to go on to pick up additional cues as to that which was unknown, and that his substitution, in the final analysis, sounds like language and retains the author's meaning.

Linda, a fifth grade student, demonstrates what this model looks like in operation. Notice particularly that reading is not perceived as an exact process, but that the author's meaning has been preserved.

> is
> In summer the living has been easy for the world's largest
> animals
> carnivorous land mammals, the big brown bears of Kodiak Islands.[3]

KEY
carnivorous = omission
animals
mammals substitution

CONCLUSION

This paper has attempted to capture some of the field data and thinking which led us to the formulation of the hypothesis that both the teaching and learning of reading are theoretically based. Readers interested in pursuing this hypothesis should contact the authors as new research instruments and procedures are currently being field tested.

It is our belief that the findings reported in this paper merit widespread exploration and have much utility for the profession. An experience we had while involved in teacher preparation follows and makes this point most vividly.

We called one teacher, whom we had observed conducting a directed teaching lesson in which children were presented vocabulary, set a purpose for reading, read the selection, and did follow-up skill work, to check out a date when our classes might observe her teach. To our surprise she announced that we could bring our students on the date we wished, but that she would be doing a "language experience" lesson. We agreed to come with the understanding that she would talk about how she normally conducted reading with our students at the end of the hour. When we arrived she did teach a language experience lesson to the children. Together the class composed a story covering a class trip to the zoo. The story read:

We went to the zoo.
We saw lots of animals.
We saw a monkey.
We saw a tiger.
We saw a duck.
We had lots of fun.

It was fascinating watching her use this approach. No matter what the children actually said, she transformed it into the type of sentence shown in the story. When the class finished the story, the teacher framed the letters "we" and asked children to identify the word and find the same word someplace else in the story. She followed the same procedure with the word "saw." With the word "A" she said, "This is a sight word. Who can find this same word someplace else in the story?" While this teacher might have "changed reading approaches," because she had not changed theoretical orientations, what she was doing in the classroom remained, in effect, unaltered.

From this experience and others like it, we have come to believe that looking at reading instruction in terms of theoretical orientation is a more cogent, insightful, and accurate one than looking at reading instruction in terms of reading approaches. In short, the variable we have identified looks hopeful. We would encourage you to explore it with us.

REFERENCES

Aaron, I. E., et al. *Scott Foresman reading systems.* Glenview, Ill.: Scott, Foresman, 1971.

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. *Framework for comprehending discourse* (Tech. Report No. 12). Urbana, Ill.: Laboratory for Cognitive Studies in Education, University of Illinois, 1976.

McCracken, G., & Walcutt, C. C. *Basic reading: Pre-primer and primer* (Teacher's ed.). Philadelphia: J. B. Lippincott, 1963.

Martin, B., & Brogan, P. *Sounds of language program: Sounds of jubilee and sounds of freedomring* (Teacher's ed.). New York: Holt, Rinehart and Winston, 1972.

Robinson, H. M., Monroe, M., & Artley, A. S. *The new basic readers: Fun with our family and fun wherever we are* (Teacher's ed.). Chicago: Scott, Foresman, 1962.

Smith, E. B., Goodman, K. S., & Meredith, R. *Language and thinking in school* (2nd ed.). New York: Holt, Rinehart and Winston, 1976.

Van Allen, R. *Language experiences in reading (LEIR).* Chicago: Encyclopedia Britannica Educational Corporation, 1974.

REFERENCES TO MATERIALS READ BY PUPILS

[1] "The Clever Turtle." Level 6, Book B, Scott Foresman Reading Systems, 1971.

[2] Ibid.

[3] Text and story authorship unknown.

SHIRLEY B. MERLIN
James Madison University

JAMES L. LAFFEY
James Madison University

NOEL E. BOWLING
West Virginia College of Graduate Studies

DEVELOPING AND EVALUATING A CRITERION REFERENCED TEST FOR PRESERVICE READING EDUCATION STUDENTS

Recently there has been an interest in developing more precise instruments to assess student achievement in reading education courses and programs. Artley and Hardin (1971) developed and field tested the *Inventory of Teacher Knowledge of Reading*. Koenke (1976), at the annual meeting of the College Reading Association, reported on the development of an *Inventory of Knowledge of Reading* for junior and senior high school teachers. Test development in both instances appeared to be related to the acquisition of general knowledge about reading. In reviewing these tests it was noted that there is insufficient descriptive information relative to the technical aspects of test development. Technical test development procedures have been recognized as being rather complex and, consequently, are often not undertaken or made available in reports.

The purpose of this paper is to describe the development of *The Reading Competency Survey Test* which was devised as part of the Right-To-Read Preservice Reading Education Program at Madison College. The instrument was designed to assist in the evaluation of students in an elementary reading methods course. The project staff found it necessary to construct a test in order to relate the program's instructional goals and objectives directly to the assessment of student achievement.

REVIEW OF PUBLISHED TESTS

Before embarking upon the construction of a tailor made instrument, several published tests were reviewed. Two tests were found to be inadequate for the purposes of the program. The first instrument (Durkin, 1964) measured only one aspect of reading: phonics. The second test, the *ETS Reading Specialist Test* (1969-70), was too advanced for preservice students. The test was designed to evaluate reading specialists with advanced work in reading. The third instrument, the Artley-Hardin *Inventory of Teacher Knowledge of Reading*, was given serious consideration since it provided a broad assessment of teacher knowledge in reading. In addition, the validity and reliability of the instrument have been established (Kingston, Brosier, and Hsu, 1975; Koenke, 1976).

In order to determine the feasibility of using all or part of the Artley-Hardin test, the Inventory was examined to decide whether (1) the content of each item was suitable for assessing objectives related to the course, and (2) the items were reasonable; that is, at the appropriate level of difficulty for preservice students.

The results of the analysis revealed that an inadequate number, 52 of the 95 items, were appropriate for assessment of the project objectives.

At this juncture, it was agreed that the evaluation process could best be accomplished by devising an instrument to match the instructional objectives in the program. A project test development team was formed. The team first established specific test development procedures. Then, the team reviewed and discussed the characteristics and values of criterion referenced tests and compared them to standardized tests (Popham, 1971; Bloom, Hastings & Madaus, 1971). The important differences between criterion and norm referenced tests relative to the discussion are given below.

CRITERION REFERENCED TESTS AND NORM REFERENCED TESTS

Criterion referenced tests are usually related to specific instructional goals or objectives. Thus, they are more useful for assessment purposes in education than norm referenced tests which are more appropriate for the counseling function in education.

Another value of criterion referenced tests relating to the assessment function is that criterion referenced tests are more useful for diagnosis since the items are written so they relate to specific instructional objectives. In addition, the goal of criterion referenced tests is to provide information related to individual performance while the goal or intent of norm referenced tests is to provide comparative data related to group performance.

A factor which is of considerable importance to norm-referenced tests and is of lesser importance to criterion referenced tests is item discrimination. Since the fundamental goal of norm referenced tests is to describe behavior in relation to the performance of others of the same age or grade, it is critical to such tests that the items discriminate well between high scoring and low scoring individuals. This can be done by writing items which maximize test variance. These types of items are more often answered correctly by high scoring individuals than by low scoring individuals. For this reason, item discrimination, which falls in the middle ranges (i.e., .40 to .60) of difficulty, is essential. In some instances a broader range of item discrimination (i.e., .20 to .80) is acceptable (Bloom et al., 1971). For criterion referenced tests, this is not so important. The ranges can be even further extended.

In considering the question of validity, the developer must keep in mind that the goal of criterion-referenced tests is to describe behavior as accurately as possible in relation to carefully prescribed instructional objectives established before the implementation of measurement procedures. This is accomplished most often by carefully writing items which relate directly, on a judgmental basis, to a set of standards or criteria for which the test iems have been designed. In this instance, content validity is fundamental to the success of the measurement activity. Womer (1975), in fact, emphasizes that writers of criterion-referenced instruments should give first priority to content validity relating the test items to specific objectives.

In order to carefully attend to the content validity of the test, a table of specifications was developed. The test items were based on a list of general competencies and specific instructional objectives which related to the instructional modules being written for the developmental reading course. Outlines of the general competencies and instructional objectives were reviewed and approved by two advisory committees consisting of reading and education faculty members, graduate students, and principals and reading teachers from the public schools. The faculty groups concurred that the competencies and instructional objectives were taught by the majority of people teaching reading and language arts courses, and the public school educators indicated that the general competencies and instructional objectives were important for training preservice teachers to teach reading in the classroom.

SPECIFIC TEST CONSTRUCTION PROCEDURES

An important step in the construction of the test was the development of a table of specifications through the use of a set of instructional objectives. This table then became the chief decision-making tool for test item development. For each area of the table of specifications a number of test items were written. The areas which were considered most important in terms of content and course coverage were listed as priority areas in which a higher number of test items were to be prepared. The topical areas included in the test were: The Reading Process, Readiness, Sight Words, Word Attack, Context Clues, Phonics, Structural Analysis, Comprehension, Study Skills, Oral and Silent Reading, Lesson Plans, Behavioral Objectives, Content Areas, Cultural Differences, Informal Classroom Diagnosis, and Approaches to Reading.

Once the test item writing was underway, the test team considered the more technical problems of test development including determining the acceptability level for both the index of discrimination and the index of difficulty. In addition, the test team had to determine how it was going to deal with the problems of validity and reliability.

The problem of validity was approached by dealing with the test items in terms of a set of logical procedures. Each test item was analyzed (1) to determine if it related to the set of instructional objectives as it was purported to and (2) to determine if the logic and language of the item was reasonable. Through these procedures the content or face validity of the test was established.

To deal with the critical task of setting an acceptance level for the index of discrimination and the index of difficulty, the test team decided to delay any decision about these levels until the test was tried out with different populations of students enrolled in the introductory reading course. The test was administered to over two hundred students enrolled in a beginning reading education course in the spring of 1974-75. The acceptance level for the index of difficulty was set at a range between 10 and 90 percent and the acceptance level for the index of discrimination was established at a range between .10 to .90.

The test was administered again in the fall semester of 1974-1975. Following this administration, the test items were reviewed by the test team using the criteria noted above.

After reviewing the test items using the established criterion, it was determined that since many of the items were logically consistent and related to specific instructional objectives, the statistical acceptance levels established for the indices of discrimination and difficulty needed to be changed. Consequently, the original statistical boundaries for the index of difficulty were extended to include all items between 6 and 95 percent and the index of discrimination were extended to include items falling between .06 and .95. In addition, items in the ranges beyond original acceptance levels which at first did not meet the specifications established by the test writing team were rewritten after the common-sense criteria of logic, language, and content validity were applied to them.

Detailed information related to the index of difficulty and index of discrimination is described in the section below.

STATISTICAL ANALYSIS

The results of the statistical analysis to determine the index of difficulty revealed that only 6 out of 125 items fell beyond the extreme limits of acceptability, 0-5 or 96-99 percent. There were 119 or 95 percent of the total items which fell in the acceptable ranges established by the test construction team. It is important to note that the index of difficulty reflects the percentage of test takers who pass an item.

An analysis of the results of computations on the 125 items to determine the indices of discrimination indicated that 93 of the 125 items met the criteria established by the test team. That is to say, 93 items fell within the range of discrimination established, .06 to .95. Only 32 items failed to fall within these ranges. All 32 fell below the discrimination level of .05. Of these 32 items, 15 items had a negative index of discrimination. This means that on these test items, more test takers from the lower group got the item correct than test takers from the upper group. More importantly, however, the negative value of these items suggests that the items were working in the opposite direction of the total test. While low positive discriminations are acceptable, the negative items need to be revised or eliminated.

In order to establish the reliability of the test, the test authors decided to use two measures of reliability. To determine the internal consistency of the test, an odd-even analysis of the test items over 162 observations was made. The computed r was .7479. According to Bruning and Kintz (1968), .70 is considered an acceptable level of reliability for a split-half coefficient of internal consistency.

To determine how reliable the test was under test-retest conditions, a small population of 20 students was tested in December, 1975, at the end of the introductory reading survey course and again in early January, 1976. A Pearson product-moment correlation was computed on the two sets of test scores. The obtained correlation ($r = .83$) indicates that the test appears to be highly reliable.

SUMMARY AND CONCLUSIONS

This report describes the development of *The Reading Competency Survey Test* designed to assess student growth in a beginning reading survey course. Before undertaking the construction of a tailor-made instrument the test development team first reviewed tests currently available in reading education. When this approach was not fruitful, the test team analyzed the strengths of criterion referenced tests and compared them to norm referenced tests. The group adopted specific test development procedures. This included the utilization of competency and instructional objectives into a table of specifications, and determining validity and reliability criteria for the test. Finally, this report describes the results of the statistical analysis of the test items and reliability data.

In conclusion, as a result of the undertaking, the authors note that while significant progress was made in constructing *The Reading Competency Survey Test,* they also recognize the need to continue to evaluate and revise specific items and particular domains of the test. Thus, it is apparent that the development of an instrument for assessing students' attainment of competency in reading is not a casual short-term project but represents a long term on-going commitment of time and energy.

REFERENCES

Artley, A. S., & Hardin, V. B. *Inventory of teacher knowledge of reading.* Columbia, Missouri: Lucas Brothers Publishers, 1971.

Bloom, B. S., Hastings, T. J., & Madaus, G. F. *Handbook on formative and summative evaluation of student learning.* New York: McGraw Hill, 1971.

Bruning, J. L., & Kintz, L. *Computational handbook of statistics.* Glenview, Illinois: Scott, Foresman Co., 1968.

Durkin, D., & Meshover, L. *Phonics knowledge survey.* New York: Columbia Teacher's College Press, 1964.

Educational Testing Service. *National teachers examinations: Reading specialist — elementary school.* Princeton, New Jersey: Educational Testing Service, 1969-70.

Kingston, A. J., Brosier, G. F., & Hsu, Y. The inventory of teacher knowledge of reading — a Validation. *The Reading Teacher,* 1975, *29,* 133-136.

Koenke, K. Ascertaining knowledge of reading with the Artley-Hardin inventory. *Reflections and Investigations on Reading.* Clemson, South Carolina: The National Reading Conference, Inc., 1976, 141-146.

Koenke, K. *Development of an inventory of knowledge of reading for middle and senior high school teachers.* Paper presented at the annual meeting of the College Reading Association, Miami Beach, Florida, October, 1976.

Popham, J. W. (Ed.). *Criterion-referenced measurement.* Englewood Cliffs, New Jersey: Educational Technology Publications, 1971.

Womer, B. What about criterion referenced measurement? In W. E. Blanton, R. Farr, & J. Tuinman (Eds.), *Measuring reading performance.* Newark, Delaware: International Reading Association, 1975.

ROBERT L. AARON
University of Georgia

ALFREDO STOKES
ETTA BATES
Dougherty County Schools, Georgia

THE RELATIONSHIP BETWEEN TEACHER ASSESSMENT OF PUPIL MASTERY AND STANDARDIZED READING ACHIEVEMENT SCORES

All the classroom teacher of reading needs is a good, well-coordinated system of worksheets, alternate teaching sources, criterion-referenced skill tests, progress charting, and proper basal placement to have a good pupil achievement design.

That idea *is* quite logical and is a thesis generally acceptable to reading specialists and to reading teacher educators. However, clear-cut research data to support such a broad generalization is lacking. Otto (1973), reported on a study of primary pupils' reading achievement gains when one group had the benefit of a classroom management system and the other group continued in the traditional basal reader centered program. Differences in reading achievement were non-significant, but "trends favored the experimental group," according to the study.

Quilling (1970), analyzed the Gates-McGinitie standardized reading test results for the first grades in two schools using a management system and found that "the mean grade equivalents for both vocabulary and comprehension were improved" over past test results. However, the results obtained did not appear to be significant beyond an acceptable chance level.

Gordon (1971), summarized the available findings on compensatory education and found general support for the structured, organized curriculum concept. He indicated that the disadvantaged pupil profits most from the tightly structured, programmatic approach that contains much individualization and immediate feedback.

Clearly, the research available on the use of highly structured "systems" in the primary grades is inadequate from the standpoint of evidencing any correlations between teacher-efficiency in using the components and pupil achievement gains. This research project sought to determine the relationship between teacher grade placement of pupils on two such components — mastery charts of sight words and basal reader reading level — and the pupils' standardized reading achievement scores in word study and paragraph meaning.

METHOD

Subjects

Sixty-nine first grade teachers of 1823 pupils in twenty-one different schools were included in the study. The schools were divided into five socioeconomic categories based on the percentage of pupils in each school considered by federal income guidelines to be economically deprived. This was done to allow a comparison of schools having more equivalent pupil populations, since intelligence tests were not available by which to equate pupil performance. The schools ranged from having less than 1% of the pupils considered to be deprived to almost 60% deprivation. Socioeconomic intervals used were as follows:

% deprivation

Very High = 1 to 5%

High = 6 to 10%

Average = 11 to 30%

Low = 31 to 40%

Very Low = 41 to 60%

While the schools listed in the two intervals identified by the parameters of 1 to 10 percent were very continuously placed, those in the last three intervals were much more widely spaced. The third and fifth intervals contained a wider poverty range, but the divisions were made at points where a significant difference existed between the last school in an interval and the first school in the next interval.

Procedure

Pupil reading achievement scores were compared with their teacher's placement of the pupils on the sight word and basal reader charts. The publisher's grade designations were used as the placement intervals and pupil test scores were correlated with these intervals for both the sight word data and the basal reader data, in the following way:

Basal and Sight Vocabulary Intervals	Grade Placements on Standardized Tests
1.0	1.0 - 1.1 - 1.2
1.3	1.3 - 1.4
1.5	1.5 - 1.6 - 1.7
1.8	1.8 thru 2.4
2.5	2.5 thru 3.4
3.5	3.5 thru 4.4

This design was used in order to consider the tendency of the standardized test to overestimate the instructional reading level of the pupil.

RESULTS

The results for all twenty-one schools in socioeconomic categories 1 to 5 were summarized under the headings of Underplacement, Correct Placement, and Overplacement. Underplacement meant the pupil was placed from one to three book or sight word intervals below his/her score on the standardized word study and paragraph reading subtests. Correct Placement meant the pupil's basal and/or sight word placement matched the standardized word study and paragraph reading subtests. Overplacement meant the pupil was placed from one to three book or sight word intervals above his/her score on the standardized word study and paragraph reading subtests.

A summary of the percentage of pupils underplaced, correctly placed, and overplaced in both vocabulary and comprehension achievement indicated the following: markedly higher percentages of pupils were underplaced by the top two socioeconomic categories and, conversely, overplaced by the bottom three socioeconomic categories.

A two-by-two chi^2 analysis was performed comparing each socioeconomic group with every other one.

On sight word list mastery the two highest groups (1-10% pupil poverty) underplaced pupils to a significantly greater degree ($p < .001$) than the bottom three categories did when these placements were compared with their standardized test scores.

In the correct placement category the bottom two socioeconomic categories (30-60% pupil poverty) were significantly better ($p < .001$) at properly placing pupils than was the top category (1-5% pupil poverty).

Overplacement is a distinctly negative behavior since it involves placing pupils significantly higher in sight vocabulary mastery than they scored on the standardized vocabulary subtest. The top two socioeconomic groups were significantly better at avoiding this negative behavior since they overplaced considerably fewer pupils than the other three socioeconomic categories.

Table 1

Vocabulary

Placement of pupils by percentage

Socioeconomic Group	Underplaced		Correctly Placed		Overplaced	
	N	%	N	%	N	%
1 - 5%	316	65	139	28	35	07
6 - 10%	245	57	155	36	30	07
11 - 30%	108	35	108	35	89	30
31 - 40%	136	38	150	42	74	20
41 - 60%	84	36	96	41	56	23

In the comprehension score-basal reader placement comparison the top socioeconomic category was significantly different ($p < .001$) from only the middle group in underplacement of pupils.

There was no significant difference among the five socioeconomic groups in correct placement of pupils in a basal reader.

The top two socioeconomic groups were significantly different from the bottom three groups in that they placed considerably less (one-half the bottom groups' rate) pupils in too high a basal reader for instructional purposes.

Table 2

Comprehension

Placement of pupils by percentage

Socioeconomic Group	Underplaced		Correctly Placed		Overplaced	
	N	%	N	%	N	%
1 - 5%	221	45	211	43	58	12
6 - 10%	180	41	206	47	49	12
11 - 30%	68	22	161	53	76	25
31 - 40%	109	30	146	41	105	29
41 - 60%	68	29	105	45	63	26

In order to determine if the significantly better placement ability of the teachers in the high socioeconomic schools was a function of the more able school population, the following analysis was undertaken.

From the total population of sixty-nine teachers, nine each with the highest and lowest pupil achievement scores, determined by covarying out the differential effects of readiness scores, were compared for their ability to properly place their pupils on sight vocabulary and basal reader mastery levels. As before, the teachers were described as having under, correctly, and overplaced their pupils.

In the vocabulary placement analysis sample five of the nine teachers with the highest achievement scores were from schools in the bottom two socioeconomic categories, while only two were from the top socioeconomic schools. Among the teachers whose pupils had the lowest achievement scores were one each from schools in the highest and lowest income categories.

The high comprehension score category contained three teachers from the lowest income schools and only two from the highest income schools. The lowest achievement group contained two teachers each from the highest and lowest income schools.

In both vocabulary and comprehension there were significant differences ($p < .001$) in pupil placement between the teachers who got the highest and lowest achievement scores. The group with the highest pupil scores was different from the group with the lowest scores in that they overplaced significantly fewer pupils and underplaced significantly more.

Table 3

Analysis of pupil placement by teachers of highest and lowest achieving pupils

Vocabulary

Pupil Placement	Underplacement			Correct Placement			Overplacement		
	N	%	X^2	N	%	X^2	N	%	X^2
Teacher Category									
High pupil ach.	122	57	*50.79	70	32	8.73	24	11	*24.61
Low pupil ach.	63	24		118	46		77	30	

Comprehension

Pupil Placement	Underplacement			Correct Placement			Overplacement		
	N	%	X^2	N	%	X^2	N	%	X^2
Teacher Category									
High pupil ach.	166	52	*36.47	76	34	2.89	30	14	*24.65
Low pupil ach.	57	25		97	42		77	33	

*$p < .001$

DISCUSSION

Apparently the system of charting pupil sight vocabulary and basal reader mastery progress did little to alter the ability of the teachers with low achieving and low socioeconomic pupils to more properly place and instruct the pupils on levels where optimum learning could occur.

The analysis of the highest and lowest achieving classes supports the importance of underestimating pupil achievement levels. The highest achieving classes' mean gain scores were significantly above the mean gain score of the total group. They underestimated more of their pupils' levels than did the low achieving groups' teachers who usually overestimated their pupils' achievement levels.

One implication of the data is that the progress charting scheme, instead of complementing the basal reader sequence, may tend to subvert the teachers' classroom developed ability to evaluate the pupils' proper basal reading level. This may be especially true if the system too strongly emphasizes "skill" mastery, as this one did, as a major criterion for determining reading achievement level. Although, comparisons between schools with and without mastery systems would be needed to verify this speculation.

To an efficiency expert entering the room of the high achievement teachers and studying their pupil placement charts they would appear to be less competent than the low achievement teachers. They consistently rated their pupils lower than did the low achievement teachers; yet, they came out with the highest achievement scores of the group of 69 teachers. Thus, they must be slowing down the pace demanded by the skills management system. A logical assumption is that they are demanding more practice, more mastery evidence, and are giving more positive feedback to their pupils about their successes. There is classroom observation evidence which suggests that these teachers require more sustained performance of the word recognition skills in the context of the basal reader than do the low achievement teachers, who tend to rely more faithfully on the results from the skill mastery tests.

While the rigorous progress charting concept may someday replace the "mastery of the basal" design it currently appears to suffer from the same variations in *quality of use* as does the basal reader. Whatever the classroom structuring concept may be it still seems to work better if the dog (teacher) "still wags the tail."

REFERENCES

Gordon, E. W. *Utilizing available information from compensatory education and surveys, final report.* Washington, D. C.: U.S. Office of Education, 1971.

Otto, W. & Askov, E. *The Wisconsin design for reading skill development. Rationale and guidelines.* Minneapolis: National Computer Systems, 1973.

Quilling, M., & Graper, N. *A comparison of school performance in reading. Technical monograph.* Madison: Wisconsin Research and Development Center for Cognitive Learning, 1970.

C2

MARY M. DUPUIS
EUNICE N. ASKOV
Pennsylvania State University

DIAGNOSTIC TEACHING IN CONTENT AREA READING

Diagnosis is not a concept which is discussed frequently with content teachers. When diagnosis enters, it usually is construed as diagnosis of students' cognitive level, of their mastery of content material. Diagnosis of reading level has been presumed to be the responsibility of the reading teacher, or for secondary teachers, the oft-maligned elementary teacher. Our thesis in this paper is two-fold: (1) that informal diagnosis of reading levels is the responsibility of the content teacher, and content teachers can develop, administer, and evaluate informal diagnostic instruments; and (2) that content teachers can use diagnostic information, teach the necessary reading skills, and not neglect their content responsibilities.

THE RATIONALE

Needed Diagnostic Information

Content teachers need to know a great deal about the reading skill levels of their students (Shepherd, 1973). First, teachers must analyze the objectives and materials on which their courses are based to determine which skills students must be able to use in order to be successful. Then they must determine from earlier teachers, curriculum guides, etc., which skills the students have already been exposed to. Questions to be answered include: Which of these skills have the students mastered? Can students transfer skills learned in reading class (or any other class) to the content material in this class?

A great many skills are introduced in the elementary grades in the context of reading class; e.g., word attack skills, such as using roots and affixes to find the meaning of unfamiliar words. The task of the secondary content teacher is to teach consciously for the necessary transfer of this basic word attack skill, reinforcing and expanding the skill for particular use in the given content area.

For example, in roots and affixes, students enter 7th grade with some familiarity with inflectional affixes and some basic derivational affixes. The junior high science teacher must begin using specifically scientific affixes and the heavy Greek and Latin roots so common in scientific language. Diagnostic questions include:

1). Can students identify and apply (two different levels), inflectional affixes in non-scientific words?

2). Can students take basic scientific terms apart into component parts?

3). Can students build new words from groups of roots and affixes and identify probable meanings?

After the basic skill is transferred and reinforced, the science teacher can introduce more elaborate root and affix work, combining basic sound pattern rules (again, familiar from elementary school) and the word-building process (consider the movement from *microscope* to *microscopic* to *microscopy*). The complexity of the process of vocabulary building is easy for reading teachers to see, but content teachers are often unaware of it.

This latter set of word attack skills, synthesizing several basic skills in the approach to unfamiliar words, represents the second major diagnostic problem for content teachers. Once students' mastery of skills taught earlier is ascertained, content teachers must assess their mastery of developing skills, ones not assumed to have been mastered earlier. On such skills as these, students will need direct instruction, not merely transfer and reinforcement. These skills will need sequential work, carefully planned and systematically introduced. Examples of developing skills include higher level problem-solving and searching skills; drawing abstract conclusions, like "theme" of literature; appreciating figurative language.

A third diagnostic problem for content teachers at all grade levels is to match the reading level of texts and other pieces of reading to the reading abilities of the students. This requires teachers to assess the difficulty of texts. It also requires them to assess the reading levels of their students. With this information, teachers can find reading materials appropriate for the students or, if that is not possible, they can adjust their teaching and the students' uses of the material to account for differences. Most content teachers are not trained either to assess materials or to assess student reading levels. Nor are they trained in adjusting their teaching or their students' reading. Nor are they familiar with alternative and supplemental materials.

Using Diagnostic Information

The Content Area Reading Project proposed to help junior high content teachers develop the skills necessary to answer these and other questions and problems in incorporating the teaching of reading in their special content. The project aimed to have its teacher participants learn about reading techniques, develop usable models, and evaluate their use in the classroom.

The long-range goal is the synthesis of necessary reading skill instruction with regular content instruction. Selection of content goals is the primary consideration, with the students' mastery of these content goals the critical outcome. However, since such mastery will not occur without adequate reading skills, these content goals must be coupled with reading skill instruction for those students who need it. This requires the application of the diagnostic information discussed earlier integrating the reading objectives with those in the content area.

THE DECISION MODEL FOR DIAGNOSTIC TEACHING

The key concept in the project's approach to content teachers is the Decision Model for Diagnostic Teaching (Cartwright, Cartwright & Yssledyke, 1973). This

64

model, based on analysis of each student's potential, requires professional decisions at specific points in the teaching situation. The original decision model was designed for teachers of special education students in regular elementary classes. The model works well in any class in which individualized attention can be given to a student.

Realistically, individual attention to each student is not possible for secondary content teachers, faced with 150 or so students each day, for discrete periods of time, with little flexibility in space. A more reasonable model for such teachers is the adaptation of the Decision Model given in Figure 1. This adapted model differs from the original in several important ways. However, it is essentially the same: a step-by-step planning sequence for a teacher to follow throughout the teaching-learning process.

Step one is a careful assessment of each student to identify relevant characteristics, in this case of his reading ability. The project focused on informal assessment of important reading skills by the content teacher using his own materials. Teachers developed group informal reading inventories (Shepherd, 1973; Thelen, 1976; Earle, 1976) and cloze techniques (Dupuis, 1976; Riley, 1973) and used classroom observation techniques. Case studies of individual students focused on background information and student attitudes and interests. Many teachers have done this, but for too many, this valuable diagnostic information is gathered, then never used. The rest of the model forces teachers to act on this information.

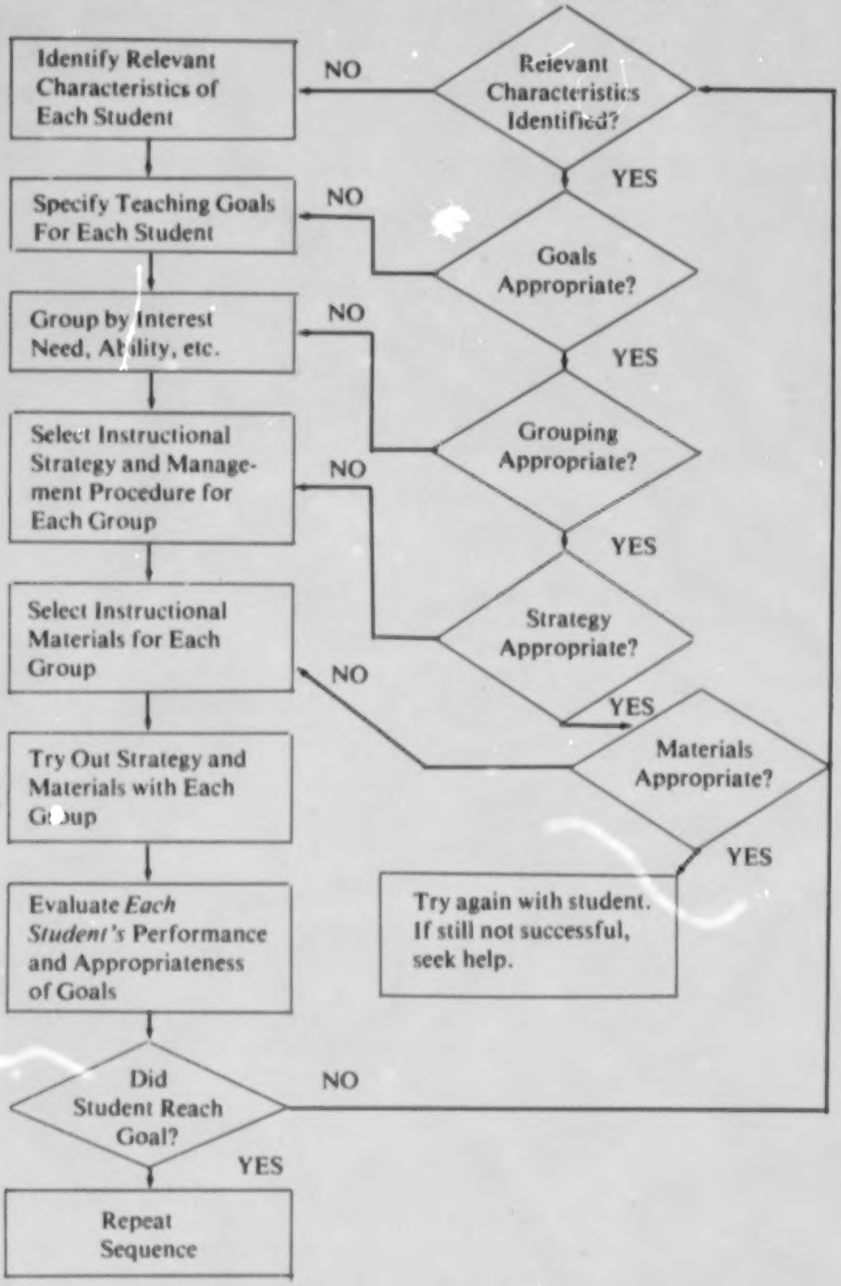
Step two asks teachers to specify teaching goals for each student. This will require writing objectives which speak to learning, both in the content area and in reading skills (Dupuis, 1973). The objective prescribed for each student may differ in which reading skill is being learned, or in what medium is being used as the source of content information (e.g., book, videotape, microfilm, magazine article, audiotape, etc.). The objective may allow choice in the method of demonstrating knowledge (e.g., oral or written, creative project or research paper), or choice in the route to gaining the skill or information (e.g., class work, learning centers or learning packets).

It seems reasonable that these student objectives be considered in groups or clusters rather than individually. That is, objectives can be grouped together in clusters of 3 or 4 which are somehow similar: the reading skill is the same, the topic is the same, the process is the same, etc. These clusters are found to be useful for several students, hence a group of students working on the same cluster of objectives.

Step three of the model is the process of grouping. The teacher makes the professional decision that certain objectives for certain students can be achieved in a group. These groups can be based on skill strengths or weaknesses, interests, special abilities; or a group of students can be matched to particular alternative reading materials. Such groups differ from typical homogenous groupings because they are re-evaluated after each sequence. They can be changed at any time: when weaknesses have been corrected, interests have changed, reading materials require different combinations.

The project trained teachers in forming different kinds of groups. The demonstration phase asked teachers to apply this procedure in a class and to

Figure 1
Decision Model for Diagnostic Teaching by Grouping


Identify Relevant Characteristics of Each Student
NO
Relevant Characteristics Identified?
YES
Specify Teaching Goals For Each Student
NO
Goals Appropriate?
YES
Group by Interest Need, Ability, etc.
NO
Grouping Appropriate?
YES
Select Instructional Strategy and Management Procedure for Each Group
NO
Strategy Appropriate?
YES
Select Instructional Materials for Each Group
NO
Materials Appropriate?
YES
Try again with student. If still not successful, seek help.
Try Out Strategy and Materials with Each Group
Evaluate *Each Student's* Performance and Appropriateness of Goals
Did Student Reach Goal?
NO
YES
Repeat Sequence

evaluate its usefulness. Staff members observed these classes and discussed the results with the teachers. In addition, project workshops demonstrated different grouping patterns within workshop sessions.

Step four of the model directs the teacher, perhaps with the students, to choose which instructional strategy to use — lecture, guided discovery, inquiry, small group discussion, independent lab work, research in the library, learning packages. Each of those strategies uses different skills, requires different kinds and amounts of reading. Each allows the teacher to direct students to use their strong skills and/or to build up their weaknesses. The classroom management procedure is dictated by the instructional choices. Will the class be in one large group? Small groups? Independent study in the classroom? Somewhere else? The teacher remains responsible for knowing what each student or group is doing. The teacher must have control over all the students at all times. During this project, content teachers began to move into grouping patterns other than the single large group. Our hypothesis is that this change can only happen slowly and in small increments. Beginning with teachers' present grouping plan, other patterns can be introduced one at a time, for short periods of time. As teachers become comfortable in different situations, as they know they can keep control of the class, they are encouraged to try additional patterns.

Step five is the one which seems out of place — selecting instructional materials. The prevailing widsom is that this is the first step — the adopted textbook is the material. The Decision Model suggests that we put the horse and cart in proper order, and make the materials serve the goals and procedures which the teacher has selected. The text may still be the appropriate material. However, vast numbers of alternative materials are available. The project demands that teachers become familiar with many different materials in their content area, assessing each in terms of its suitability in reading level, skills, and content material. An additional component at this step is learning to adapt texts for use by students who can't otherwise use them: developing study guides (Thomas & Robinson, 1973), reasoning guides (Herber, 1970), even rewriting the critical parts of the content material.

Step six directs the teacher to try out these plans with each group. Each teacher in the project tried out the techniques discussed in workshops in at least one of their classes. Their feedback is step seven — evaluation. At this point, the model shifts its emphasis from the groups formed after step two back to the individual student. Each student's individual performance must be assessed to determine whether he met the objectives set for him and his group. This means, of course, that students in the same class may be reaching different objectives. While the content objectives may be the same for all of them, the reading skill objectives may well differ.

The decisions to be made during the evaluation are included in the diamond shape on the model. If the student reached the goal set, then the sequence can be repeated. If he didn't reach the goal, the teacher looks for the reason why. At which step did the teacher's planning go awry? Suitable revision should allow the student to succeed on his next attempt. It is at this point that grouping changes if the needs of the group have changed. Indeed, based on a new set of goals for each student, the grouping patterns will necessarily change frequently. Thus, the instructional pattern described here will not lock a student into a "track" or any kind of continuing group. The teacher's professional judgment will be used after each instructional sequence to be sure that the groups are serving the best interests of the student.

EVALUATION OF THE MODEL

The Decision Model is based on a diagnostic-prescriptive rationale, useful for content teachers despite the large number of students they teach. However, the project predicted that teachers would have difficulty adopting this model and making it operational. Hence we provided a full year of work, in workshops and in teachers' classrooms, in an attempt to change teacher behavior.

The model was made operational by introducing to the workshop teachers various instructional techniques for the steps in the model. At the end of the project, the teachers were asked to indicate which of the techniques they had used, fully or in part, and whether they felt each technique was valuable or not. Sixty teachers at the three sites of the project responded. Their evaluation is given in Table 1.

Table 1

Teacher Evaluation of Instructional Techniques
In Percents

	Technique	Used	Valuable
1.	Informal Group Reading Inventory	87%	94%
2.	Cloze Test	92%	88%
3.	Assessment of Student Self Direction	71%	79%
4.	Concept of Linguistic Difference	47%	63%
5.	Unit	87%	89%
6.	Learning Package	54%	75%
7.	Learning Center	55%	71%
8.	Grouping Plans	88%	92%
9.	Non-Print Media	85%	90%
10.	Comprehension Questioning Strategies at Different Levels	92%	97%
11.	Vocabulary Exercises	97%	99%
12.	Study Skill Exercises	85%	93%
13.	Reading and Study Guides	85%	92%
14.	Alternative Reading Materials at Various Levels	80%	86%
15.	Readability	90%	88%

Step one becomes operational as *diagnostic* techniques. Each teacher developed an Informal Group Reading Inventory and a cloze test, based on reading in his own content area. Each one also worked with identifying students' self-direction, as a means of determining which students needed teacher direction and which students could work more independently.

The concept of linguistic differences was introduced to add diagnostic information. Videotapes showing language development and culture in the black urban setting of one workshop school and in the rural white setting of another workshop school were used to set up a discussion of linguistic difference and how it might affect reading. This concept was hotly debated in the urban school, where most teachers were already familiar with it and concerned about it, although they didn't agree on how it should be handled. Teachers in the rural and suburban schools were generally not familiar with the concept, nor did they become very involved with it. The relatively lower percentages for this technique (47% used it; 63% thought it was valuable) demonstrate that it didn't make as much impact as the other diagnostic tools.

Steps two through seven were approached as parts of long-term planning. Teachers were asked to develop a unit or use one that they had already developed; each unit followed these six steps in the model. A prime component of each unit was a set of objectives related to reading skills necessary to master the content objectives. Each objective was followed through to step seven, so that the grouping, instructional strategies, materials and evaluation were clearly planned.

In addition to the unit, each teacher also developed a learning package or learning center. These were usually constructed as part of the unit, containing a few of the unit objectives and used over a shorter period of time but planned as part of the full sequence of the unit.

Step three was evaluated also by the development of a set of grouping plans. Each teacher developed two grouping plans, one based on reading skill data from the informal inventory or cloze, and one based on other data, usually content-oriented.

At step four, teachers were encouraged to use non-print media (85% reported using it), comprehension questioning strategies at different levels (92% reported using it), skill exercises in vocabulary and study skills, and reading and study guides to aid comprehension. (85% or more reported using each of the skill exercises.)

Step five concerned selection of materials. The concept of readability was studied and applied to the teachers' own texts. 90% used that in selecting materials during this year and in planning for next year. 80% of the teachers used alternative reading materials at different levels.

Step six presented a problem for the teachers this year. They tried out many of the techniques we introduced, but there wasn't enough time to use them in the proper order or with sufficient organization this year. A follow-up study has been planned for next year in which the teachers will systematically implement the materials they developed this year.

Step seven, evaluation of students' performance, was an integral part of each instructional plan. Evaluation ranged from standard tests (multiple choice, true-false, essay), to creative writing and videotapes, projects and papers, and artistic responses to reading. Some teachers encouraged students to develop test items and practice exercises for other students to use. Although a variety of forms of evaluation were used, the project staff felt that more emphasis on evaluation would be useful. Therefore, an additional component of next year's follow-up will be the study of additional evaluation techniques.

It is reasonable to ask whether these reported percentages represent actual use or an exaggeration by the teachers. We have no absolute proof that the techniques were used. However, project staff members, in each school at least one day a week, reported observing teachers using the plans. Staff responded to questions on implementation and reviewed student papers and projects afterwards. Teachers' logbooks reflected the use of many techniques; they enclosed dittos, lesson plans, tests, student papers, and other evidence that they had in fact implemented the techniques singly or in sequence. We are satisfied that the majority of teachers had conscientiously attempted to implement their plans. This problem with implementation represents another goal of the follow-up study, a systematic and recorded attempt to verify the usefulness of the techniques and the entire Diagnostic Teaching Model.

The prediction that teachers would have difficulty implementing the model was borne out by the work of the sixty teachers who completed the Content Area Reading Project. However, their evaluation is positive: the model provides a reasonable framework within which a teacher can synthesize the teaching of content material and reading skills.

REFERENCES

Cartwright, G. P., Cartwright, C. A., & Yssledyke, J.E. Two decision models: Identification and diagnostic teaching of handicapped children in the regular classroom. *Psychology in the Schools,* 1973, *10,* 4-11.

Dupuis, M. M. The cloze procedure: Can it be used with literature? *Reading Improvement,* 1976, *13,* 199-203.

Dupuis, V. L. A design for accountability in curriculum evaluation. *The New Campus,* 1973, *26,* 11-16.

Earle, R. A. *Teaching reading and mathematics.* Newark, Delaware: International Reading Association, 1976.

Herber, H. L. *Teaching reading in content areas.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

Riley, P. M. *The cloze procedure — A selected annotated bibliography.* Lae, New Guinea: PNG University of Technology, 1973.

Shepherd, D. L. *Comprehensive high school reading methods.* Columbus, Ohio: Charles E. Merrill Publishing Company, 1973.

Thelen, J. *Improving reading in science.* Newark, Delaware: International Reading Association, 1976.

Thomas, E. L. & Robinson, H. A. *Improving reading in every class: A sourcebook for teachers.* Boston: Allyn and Bacon, 1976.

70

ROBERT A. PALMATIER
Northwestern State University of Louisiana

SUSAN S. STRADER
Johnson City, Tennessee

TEACHER PERFORMANCE IN ASSESSMENT OF COMPARATIVE READING DIFFICULTY OF CONTENT MATERIALS

Secondary teachers have often been advised to select easier reading texts for their poorer reading students (Marksheffel, 1966). The wisdom of this suggestion is most frequently overlooked in light of the reality that a particular class is provided but one choice of text. While recognizing that teachers may not have the opportunity, this study examines their ability to make decisions concerning the reading difficulty of text material.

If teachers are expected to follow exhortations that they select reading material of appropriate difficulty for their students, it is relevant to evaluate their ability to judge the relative difficulty of text material. The study reported here examined the ability of secondary teachers to judge and rank order the difficulty of five subject area reading selections. Attention was given to the accuracy of ranking and to differences between content area teacher groups.

METHOD

Subjects

The sample comprised two groups of teachers involved in inservice activities with the senior author. Each group included the total faculty of large high schools in Bessemer, Alabama, and Douglas County, Georgia. Each school provided approximately half of the total sample (N = 180). Thus, the subjects could be said to have been nearly equally drawn from city and suburban school systems. Although it is hazardous to generalize the findings of the study beyond the particular subjects used, it is likely that these teacher groups are representative of the general population of secondary content teachers.

For this study the teachers were considered as a total group and separately in seven content area groupings. Groups were defined by the major teaching responsibility of each teacher. The seven areas were social studies, mathematics, science, English, physical education, career education, and a miscellaneous group comprised of foreign language teachers, librarians, music teachers, and art instructors. Group sizes ranged from 11 to 44 with the physical education and career education groups being considerably smaller than any of the other groups.

Procedures

Five passages, of 100 to 150 words in length, were selected from the text *Six-Way Paragraphs* by Pauk (1974). The selections were made to represent reading difficulties ranging from seventh to eleventh grade readability levels as estimated using the Fry Graph (Fry, 1968). In addition to varying difficulties the passages were chosen to represent the five content areas of sports, mathematics, history, science, and literature. The passages were reproduced with each selection on a separate page.

Within the context of inservice workshops, subjects were presented with the task of rank ordering the passages according to their reading difficulty for students. This task was done prior to instruction on the determination of readability. Questions concerning how decisions should be made were answered with "whatever method seems best to you." After making decisions, teachers were asked to indicate their rankings on a separate form. Time was allowed for all to complete the task. Both groups finished in less than 15 minutes. Prior to announcement of actual rankings and instruction in a method for determining reading difficulty, ranking forms were collected.

Data were keypunched and subjected to computer analysis using a parametric portion of the Statistical Package for the Social Sciences (SPSS, 1973). The portion of the program used provided frequency and percentage tabulation for each group and for the total sample on each of the five passages. The program also provided Chi Square evaluation of the associations between content teacher groups and accuracy of rankings.

RESULTS

Teacher groups were generally unable to rank the five passages accurately according to reading difficulty. In no case was a single passage ranked correctly by more than 29 percent of the total sample. Four of the seven teacher groups were able to most frequently identify the correct ranking of the literature selection. Two groups achieved a greater percentage of accuracy on the sports selection while one group was correct most frequently on the math passage. The incidence of accurate designation, by separate content groups, of rank for a particular content passage ranged from 0 to 57 percent. No pattern could be recognized in the relationship between the content of passages and the subject area taught by teachers in designated groups. Social studies teachers for instance were best in accurate ranking of the literature passage, poorest in the ranking of the science selection, and only able to correctly rank the social studies selection in 14 percent of the cases.

From observation it appears that the ranking of the science passage was the most difficult task for the teachers, as only 9 percent reported the correct rank. The history passage was also ranked correctly by only a small number of the sample(15%). It is of interest to note that science teachers were not more able to accurately identify the ranking of the science passage than were other teacher groups. Prior to the experiment it was thought that there might be some relationship between the subject area taught and the ability to rank particular content passages. Chi Square values were calculated in order to assess this speculation. In only two cases was a significant association found between accuracy in ranking a specific passage and separate teacher groups. The sports and science passages were ranked by the teacher groups with a significantly differing level of accuracy.

Calculation of Cramer's *V* in these cases indicated that there was only a weak pattern of association between accuracy of rankings and separate teacher groups. Observation led to the suspicion that the polar positions of the passages (sports, easiest and science, most difficult) may have interacted to provide an unwarranted finding of significant differences.

In attempting to assess the accuracy of correct rankings by content groups extreme caution must be exercised. The fact that the two smallest content area groups achieved the best record for accurate rankings led to the speculation that the number of subjects in these groups was not sufficient to provide reliable measurement. That only two of the seven groups' total ranking accuracy deviated from a range of 18 to 22 percent added further strength to the suspicion of low reliability in rankings by small groups of physical education (25%) and career education (31%) teachers.

DISCUSSION

The experiment primarily demonstrates that the task of making decisions as to the reading difficulty of material is not one in which the high school teachers in this sample were competent. While the normalcy of this sample cannot be argued on the basis of sampling it is suspected that the teachers selected represented the broader spectrum of secondary content area teachers reasonably well. Thus, the conclusion is easily supported that methods for determining reading difficulty should be established as a high order of priority for high school teacher inservice activities. Further study of this task might be conducted to determine if even the knowledge of how to determine readability by formula is sufficient to allow for the sort of holistic "eyeball" determination of relative difficulty involved in this study. Comparison of the content teacher population with a group of reading specialists trained in the use of readability formulas might provide further information. It would also seem wise to replicate this study using graded passages within each of the content areas. Such a replication would allow evaluation of content specific teacher judgments.

REFERENCES

Fry, E. A readability formula that saves time. *Journal of Reading*, 1968, 11, 513-516; 575-578.

Marksheffel, N. *Better reading in the secondary school.* New York: Ronald Press, 1966.

Pauk, W. *Six-way paragraphs.* Providence, Rhode Island: Jamestown Press, 1974.

Statistical package for the social sciences. Evanston, Illinois: Vogelback Computing Center, Northwestern University, 1973.

ROSALIND STREICHLER
JOSEPH S. NEMETH
Bowling Green State University

COMPETENCIES FOR COLLEGE DEVELOPMENTAL READING TEACHERS

Considerable effort has been expended in designing curricula to prepare reading teachers, clinicians, consultants, and supervisors for service at the local education agency level. Many states specify certification requirements for reading personnel. In addition, reading certification requirements have been suggested by professional organizations such as the International Reading Association (Professional Standards and Ethics Committee, 1968). In contrast, insufficient curricular attention has been focused on the preparation received by college developmental reading teachers.

In an effort to better prepare the personnel needed to staff the college reading programs, concerned members of the profession have developed heuristic training programs. For example, Maxwell (1966, 1969) has reported on several training programs which were developed to prepare college developmental reading teachers. Wolfe (1966) and Price and Wolfe (1967) have offered suggestions for the preparation of specialists in teaching college developmental reading. Also a symposium on the training of college and study skills specialists conducted during the twenty-second annual meeting of the National Reading Conference resulted in the publication of several descriptive and research based papers (Nacke, 1973). A sampling of the papers presented at the symposium follows: Vavoulis and Raygor (1973) reported on a research based curriculum developed at the University of Minnesota; Maxwell (1973) reviewed skill requirements for college reading and study skills specialists; Cranney, Schenck and Hellstrom (1973) presented their concerns about initiating a training program for junior college reading teachers; and Tonges (1973) reported on the development of competency-based teacher training material in college reading.

Additionally, Blanton and Smith (1975), in their review of the research on the training of college-adult teachers of reading, indicated that there is a paucity of basic research on the curricula, competencies, organization and management of programs which purport to train college-adult reading specialists. That conclusion had been confirmed by Ahrendt (1975) who noted that little research had been done to investigate the efficacy and usefulness of the training preparation received by college developmental reading teachers.

Thus, it appears that although there is an increasing interest and awareness of the need for high quality training programs for college-adult reading specialists, many of the existing programs are heuristic and have not been developed as a consequence of research based competencies.

PURPOSE OF THE INVESTIGATION

Meanwhile, there appears to be an increase in the student population at the junior college level as well as an increase in educationally disadvantaged students entering universities under open enrollment programs. In addition, an increasing number of propietary institutions are being formed, requiring the hiring of trained specialists to staff reading programs. It seems to be apparent that every effort should be expended to improve training in this area in order to provide competent personnel to staff and guide these programs.

The Reading Center at Bowling Green State University was concerned with improving the quality of its college developmental reading program as well as improving its preparation of college developmental reading and study skills specialists. Since competency-based teacher education appeared to offer some promise for restructuring teacher education, it seemed appropriate to focus research efforts on the application of competency-based teacher education to the preparation of college developmental reading teachers. A necessary first step in such research efforts was to identify the knowledge and performance competencies needed by college developmental reading teachers.

Hence, the major purpose of this study was to identify competencies needed by teachers of college developmental reading at three levels. A modified Delphi technique was chosen as the research procedure. The Delphi technique was used successfully to provide pre-conference input for the ACTE National Invitational Conference on Redesigning Teacher Education (1973). Additionally, Judd (1972) suggested that one valuable use of the Delphi technique would be to obtain consensus among the leaders in a specific discipline.

METHOD

The procedural steps in this research were as follows:

Step 1

A jury, selected from the membership list of a professional reading organization whose main focus is college and adult reading, was invited to respond to two separate rounds of questionnaires. Sixty-five of the 97 members contacted agreed to participate in the study. Included in this group were: fifteen full professors (Education, Reading and Language Arts, Elementary Education, Psychology, and English), seven associate and assistant professors, 13 directors (Reading Centers, Reading and Study Skills Centers, Developmental Education, Adult and Adolescent Reading Centers), the Dean of a College of Education of a major university, two coordinators of doctoral and graduate programs in reading, and a research psychologist. The remaining participants are included in the following categories: visiting lecturer, adjunct professor, professor emeritus, reading specialist, coordinator of tutorial services, and assistant and associate directors of college skills centers.

Step 2

In Round I, the jury was asked to respond to 86 competencies developed by researchers from a variety of content areas in education as well as from descriptions of teaching tasks. The 86 competencies were stated behaviorally and were categorized under ten major areas: a) basic reading foundations, b) foundations of reading instruction, c) orientation to college reading, d) facilitation of instruction, e) management of learning environment, f) instructional design, g) program design, h) effectiveness of program, i) supervision, and j) professional and personal development.

Step 3

The jury was instructed to assign each of the 86 competencies to one or more levels of application: Level I (entry), Level II (advanced), and Level III (specialization). To provide a common base for assignment of these levels, criteria based upon pre-service and post graduate education were specified. The jury then rated each competency as being "critical," "important," or "desirable" at each level. Thus, multiple competency ratings at several application levels were possible. Because each of the ten major categories was open-ended, the jury could add and/or revise competencies. Forty-five members of the original panel responded to Round I. The responses were coded according to the respondent's position, title, number of years' experience in reading, number of years' experience in college developmental reading, and description of type of institution. The responses were converted by computer to a mathematical continuum which ranked the competencies at each of the three application levels in order of importance.

Step 4

Those competencies which were ranked as "desirable" were excluded from further consideration. At Level I, 25 (29%) competencies were ranked as "critical" or "important" and were retained. Fifty-two (60%) were retained at Level II and 82 (95%) were retained at Level III.

Step 5

For Round II, the researchers analyzed and refined the lists, reducing the competencies to ten at each level of application by utilizing the highest-ranked competencies on each list and eliminating overlap and redundancies. Following suggestions of several panel members, assignment of competency levels for Round II was related to specific faculty positions. The respondents were instructed that competencies selected and ranked for Level I would be those knowledge and performance competencies expected of lecturers and instructors at two and four-year institutions with a minimum of experience. It was expected that those competencies selected at this level would be refined and maintained at Levels II and III.

Competencies selected and ranked for Level II would be those competencies expected of experienced master teachers at the assistant professor and professorial levels. These competencies would be refined and maintained at Level III.

Competencies selected and ranked for Level III would be those competencies expected of assistant professors and professors directly involved in the administration and supervision of the college reading program. This category would include directors of reading centers, internship supervisors, and program coordinators who might be involved in little, if any, actual teaching of college reading. However, it was expected that all competencies selected at Levels I and II would be refined and maintained at Level III.

Step 6

The jury was instructed to rank the revised list of ten competencies at each application level from most to least important.

Step 7

Thirty participants responded to Round 2 of the questionnaire. The data were tabulated and resulted in the following list of competencies ranked in order of importance.

Level 1: Lecturers and Instructors

1. demonstrate positive attitudes toward reading
2. demonstrate an understanding of the problems that affect individual achievement
3. demonstrate the ability to accurately diagnose the needs of the students, utilizing both standardized and informal diagnostic measures
4. demonstrate an understanding of the procedures which aid in providing for positive interaction with students
5. demonstrate the techniques for developing a flexible reading approach — judging different types of materials, selecting and combining techniques
6. demonstrate the techniques for teaching specific skills in study reading — procedures for perceiving organization and structure, determining central theme, locating main points and supportive details, and locating information efficiently
7. demonstrate the techniques for teaching specific skills in reading in the student's content area textbooks — note-taking, outlining, underlining, annotating, summarizing, perceiving paragraph organizational patterns (listing, time-order, comparison-contrast, cause-effect), preparing for examinations, transferring critical reading techniques to content area texts
8. demonstrate the techniques for teaching specific skills in vocabulary enrichment — utilizing context clues, analyzing root words and affixes, adding specialized vocabulary
9. demonstrate the techniques for teaching specific skills in critical reading — questioning, inferences, relating and evaluating
10. demonstrate efficient and functional reading skills and reading habits

Level 2: Master Teachers

1. demonstrate a knowledge of the basic skills and abilities in the following categories: word recognition, comprehension, rate flexibility, study skills, and literary habits, attitudes, appreciation, and tastes
2. demonstrate a knowledge of problems that might affect individual achievement: emotional factors, physiological factors, learning styles, cognitive structuring

3. demonstrate an understanding of a diagnostic-prescriptive teaching approach
4. demonstrate the ability to evaluate the appropriateness and adequacy of all sorts of instructional materials
5. demonstrate the ability to evaluate materials for instruction on the basis of effectiveness and appropriateness to the lessons planned
6. demonstrate the techniques for developing efficient study habits — preparing functional study schedules analyzing types of questions, preparing research papers
7. demonstrate the ability to utilize a non-commercial materials/equipment approach, understanding the principles of devising original material
8. demonstrate the understandings necessary to utilize input from subject matter teachers and provide for participation of content area faculty as tutors and resource personnel
9. demonstrate a familiarity with the operational procedures in using the following materials: programmed materials, skills development books, packaged learning systems, instructional media, reference books, library books, etc.
10. demonstrate the ability to plan for and organize the facilities required for the program

Level 3: Supervisors and Coordinators

1. demonstrate the ability to establish and administer a college developmental reading program
2. demonstrate a knowledge of the theories, models, and research regarding the reading process
3. demonstrate the ability to develop implementation strategies for new programs and provide for changes in existing programs
4. demonstrate an awareness of new theories, research, and practical applications published in professional journals
5. select and provide for appropriate staff development experiences
6. establish and administer a training program for teaching assistants and/or interns
7. demonstrate competency in orienting subject matter teachers in using reading improvement techniques to improve teaching in their content areas
8. demonstrate the ability to organize and conduct workshops and inservice education
9. demonstrate the ability to interpret and evaluate criterion and norm-referenced instruments for evaluating instruction
10. demonstrate the ability to utilize and conduct research

CONCLUSION

These competencies have been identified by consensus of a panel consisting of members of a professional organization currently holding positions in major universities, colleges, and study skill centers with from one to over twenty years experience in the field. Because of the validity afforded by the judgment of such a panel, it is believed that the resulting identified competencies form a viable base upon which to develop training programs and plan for curriculum revision. In addition, the results of this study could be adapted for teachers in other areas of reading and, in fact, could be adapted by those in other content area disciplines.

REFERENCES

Ahrendt, K. M. *Community college reading programs.* Newark, Delaware: International Reading Association, 1975.

Blanton, W. E. & Smith, C. F. Jr. 1974 Review of research on college-adult reading. In G. H. McNinch and W. D. Miller (Eds.), *Reading: Convention and inquiry.* Twenty-fourth Yearbook of the National Reading Conference, 1975, 1-47.

Cranney, A. G., et al. Initiating a program for training junior college reading teachers. In P. Nacke (Ed.), *Programs and practices for college reading.* Twenty-second Yearbook of the National Reading Conference, 1973, *2,* 178-182.

Judd, R. C. Use of delphi methods in higher education. *Technological Forecasting and Social Change,* 1972, *4,* 173-186.

Maxwell, M. Training college reading specialists. *Journal of Reading* 1966, *10,* 147-155.

Maxwell, M. What the college reading teacher needs to know about reading. *Second Annual Proceedings Western College Reading Association,* 1969, 65-70.

Maxwell, M. Skill requirements for college reading and study specialists. In P. L. Nacke (Ed.), *Programs and practices for college reading.* Twenty-second Yearbook of the National Reading Conference, 1973, *2,* 159-162.

Nacke, P. (Ed.). *Programs and practices for college reading.* Twenty-second Yearbook of the National Reading Conference, 1973, *2,* 159-186.

Price, U. & Wolfe, K. Teacher preparation of the junior college reading teacher. *Junior college reading programs.* Newark, Delaware: International Reading Association, 1967, 1-7.

Professional Standards and Ethics Committee. *Roles, responsibilities and qualifications — Reading specialists.* International Reading Association, 1968.

Tonjes, M. T. Development of competency-based teacher training material in college reading. In P. L. Nacke (Ed.), *Programs and practices for college reading.* Twenty-second Yearbook of the National Reading Conference, 1973, *2,* 183-186.

Vavoulis, A. & Raygor, A. L. The training of college reading and study skills specialists — a survey of expert opinion. Training reading and study skills specialists at the University of Minnesota. In P. L. Nacke (Ed.) *Programs and practices for college reading.* Twenty-second Yearbook of the National Reading Conference, 1973, *2,* 163-171.

Wolfe, J. B. Preparation for specialists in teaching college reading. In C. A. Ketcham (Ed.), *College Reading Association Proceedings,* 1966, 25-28.

WILLIAM H. RUPLEY
Texas A&M University

STABILITY OF TEACHER EFFECT ON PUPILS' READING ACHIEVEMENT OVER A TWO YEAR PERIOD AND ITS RELATION TO INSTRUCTIONAL EMPHASES

A general conclusion of research on teacher effectiveness in reading instruction is that an important variable with respect to pupils' achievement is the teacher (Bond & Dykstra, 1967). Limited attention however, has been given to the role of the classroom teacher in producing student achievement gains in reading (Rutherford, 1971; Farr and Weintraub, 1975). Not only are studies that focus on teachers and their effect on pupils' reading achievement small in number, many of them do not investigate the process of instruction in relation to the product.

Rosenshine (1970), McNeil and Popham (1973), and Saadeh (1970) suggest that researchers are not looking at teacher effects as they produce modifications in the learner. They assert that researchers are focusing on readily available criteria and hoping that these reflect an outcome criterion. Thus, it was decided that stability of teacher effect in reading should be determined first and then instructional process variables that could account for varying teacher effect could be explored.

One possible explanation for one teacher having a greater effect on pupils' reading achievement than another could be the emphases that the effective teacher gives to the various aspects of a self-contained developmental reading program. That is, the more effective teacher could give greater emphases to comprehension, diagnosis, oral reading, word attack, and so forth. Thus, nine hypotheses were tested, eight of which had the same format: Highly effective teachers report greater emphases on (1) specific diagnosis, (2) ongoing diagnosis, (3) comprehension readiness, (4) comprehension post reading, (5) comprehension above the literal level, (6) reading application, (7) structured reading activities, and (8) oral reading than do effective teachers and less effective teachers. The ninth hypotheses was: There is no stability between teacher's effect on pupils' reading achievement over a two year period.

METHOD

In the spring of 1974, all third and sixth grade teachers teaching reading in a self-contained classroom in the Fort Wayne Community School System, Fort Wayne, Indiana, were identified as highly effective, effective or less effective teachers of reading through the use of a least squares prediction line (Glass & Stanley, 1970). All of these teachers used a basal approach as their primary means of reading instruction. Two prediction lines were generated, one for third grade classes and one for sixth grade classes. Both prediction lines were based on class mean IQ scores determined by the Otis Lennon Mental Ability Test, Form J, administered in the fall of 1974 and the mean class reading achievement score determined by the SRA Achievement Series, administered in the spring of 1974. Individual classes were plotted in relation to the prediction line. A teacher whose class reading achievement mean was greater than one-half a standard error of estimate above the line was deemed highly effective, a teacher whose class mean fell within plus or minus one-half a standard error of estimate of the prediction line was deemed effective, and a teacher whose class reading mean fell more than one-half a standard error of estimate below the prediction line was deemed less effective.

In the spring of 1976 identical data were gathered on all third and sixth grade teachers. These data were used to generate a 1976 prediction line for sixth grade teachers. Again, all mean class reading achievement scores of third and sixth grade teachers were plotted in relation to the appropriate grade level prediction line. Identical criteria were used to identify the teachers as highly effective, effective, or less effective teachers of reading.

The 1976 sample of teachers was administered the 1976 edition of the Survey of Teacher Emphases in Reading Instruction (STERI). The STERI was developed by the investigator to assess the amount of instructional emphases given to eight components of a self-contained developmental reading program over a typical six weeks period. The subcategories included (1) ongoing diagnosis, (2) specific diagnosis, (3) comprehension readiness, (4) comprehension following reading, (5) comprehension above the literal level, (6) oral reading, (7) structured reading activities, and (8) reading application. There were 35 items on the questionnaire and the teacher responded by selecting a number which indicated the degree of emphasis given to the statement. For example, "Students are made aware that oral expression can be represented in written form." 1 (never emphasized) 2 (emphasized 1 to 3 times) 3 (emphasized 4 to 6 times) 4 (emphasized 7 to 9 times) 5 (emphasized 10 or more times).

The 1976 edition of the STERI evolved from the 1974 edition. A factor analysis of the 1974 STERI resulted in the identification of the eight subcategories. Items with a .30 or greater relationship to each other were retained and classified into instructional categories. Chronbach's alpha (1951) for the 1976 edition of the STERI was .88 and reliability coefficients for subcategories ranged from .42 to .68.

RESULTS

To determine stability of teacher effect on pupils' reading achievement over a two year period a product moment correlation was used. A random sample of 25 teachers was drawn from third grade teachers (N = 61) and a random sample of 25 teachers was drawn from sixth grade teachers (N = 101). Each teacher's 1974 standard deviation from the predicted Y variable (mean class reading achievement) was correlated with the 1976 standard deviation from the predicted Y variable. The resulting correlation ($r = 95$) was significant at the .01 level. Thus, the hypothesis that there is no stability between teacher's effect on pupils' reading achievement over a two year period was rejected.

An analysis of variance was performed to test differences in mean emphases reported by the teachers identified as highly effective, effective, and less effective both in 1974 and 1976 ($N = 41$) on the eight subcategories of the STERI. An alpha level of .15 was established apriori to data analysis.[1]

Significant differences were observed between the emphases reported by highly effective, effective, and less effective teachers of reading for the subcategories of ongoing diagnosis ($p < .09$), comprehension following reading ($p < .13$), structured reading activities ($p < .02$), and reading application ($p < .14$).

In addition, an analysis of variance was conducted to test differences in mean emphases reported for teachers who from 1974 to 1976 shifted from effective to highly effective versus those who shifted from effective to less effective (total $N = 21$) on the eight subcategories of the STERI. The level of significance was again set at .15. Significant differences were found between teachers who shifted upward and teachers who shifted downward for the subcategories ongoing diagnosis ($p < .07$), comprehension readiness ($p < .11$), comprehension following reading ($p < .09$), structured reading activities ($p < .13$), and reading application ($p < .04$).

Finally, an analysis of variance was performed to test differences in mean emphases reported by teachers identified as highly effective, effective, and less effective teachers of reading for 1976 on the eight subcategories of the STERI. The alpha level of .15 was retained for this analysis. Significant differences were noted for ongoing diagnosis ($p < .002$), comprehension readiness ($p < .15$), comprehension following reading ($p < .01$), comprehension beyond the literal level ($p < .08$), oral reading ($p < .01$), structured reading activities ($p < .01$), and reading application ($p < .01$).

DISCUSSION

The means for all three analyses favored the highly effective teachers. In addition, the means for the effective reading teachers were higher for the significant subcategories then the less effective reading teachers when highly effective, effective, and less effective teacher's emphases were compared.

The following hypotheses were supported by the data across all three analyses:

1. Highly effective teachers of reading report a greater emphases on ongoing diagnosis than do effective and less effective teachers of reading.
2. Highly effective teachers of reading report a greater emphases on comprehension following reading than do effective and less effective

---

[1] Although this alpha level deviates from those typically cited in psychological and educational research, i.e., .01 or .05, precedence for this decision can be found in the literature. Hays (1973) suggests that "one could (and perhaps should) make the test more powerful by setting the value of alpha at .10, .20, or more" (p. 360). Further, Rupley (in press) indicates, "the historical concept of setting alpha at .05 or less may be inappropriate for research dealing with effective reading instruction. The data base which is presently available for looking at effective reading instruction is minute when compared with the plethora of reading research being conducted. It might be better to say with a seventy percent degree of certainty that effective reading teachers do certain things in their reading instruction, than to say with a ninety-five percent degree of certainty that the characteristics of teacher effectiveness in reading instruction have not been identified" (p. 7).

teachers of reading.

3. Highly effective teachers of reading report a greater emphases on structured reading activities than do effective and less effective teachers of reading.
4. Highly effective teachers of reading report a greater emphases on reading application than do effective and less effective teachers of reading.
5. Teacher effect on pupils' reading achievement is stable over a two year period.

Four of the hypothesis advanced to account for teacher variance in being "successful" in reading instruction were supported. In addition, it was determined that teachers' effect on pupils' level of reading achievement is stable over a two year period.

The importance of teacher emphases in the areas of ongoing diagnosis, comprehension following reading, structured reading activities and reading application becomes more credible due to the fact that across three analyses these four variables significantly favored the highly effective teachers. The teachers who were highly effective in both 1974 and 1976 reported greater instructional emphases in these four areas than the effective and less effective teachers in both 1974 and 1976. Also, those teachers who shifted upward to highly effective in 1976 (from 1974) reported greater instructional emphases in these four areas than those teachers who shifted downward to less effective in 1976 (from 1974). Furthermore, ongoing diagnosis, comprehension following reading, structured reading activities, and reading application significantly favored the highly effective teachers for analysis of the 1976 data.

Future research aimed at determining teacher effect on pupils' reading achievement should seriously consider the instructional emphases provided by the teacher in the areas of ongoing diagnosis, comprehension following reading, structured reading activities, and reading application. Not only do these areas warrant further investigation, but future research should focus on the process variables the teacher applies in these four areas. That is, what or how do the highly effective teachers deal with the four areas in relation to effective and less effective teachers of reading? This task would require data gathering in the actual classroom setting over an extended time period.

REFERENCES

Bond, G. & Dykstra, R. The cooperative research program in first grade reading instruction. *Reading Research Quarterly,* 1967, *2,* 26-142.

Cronbach, L. Coefficient Alpha and the internal structure of tests. *Psychometrics,* 1951, *16,* 297-334.

Farr, R. & Weintraub, S. Editorial: Methodolical incarceration. *Reading Research Quarterly,* 1975, *10,* ii-iv.

Hays, W. *Statistics for the social sciences.* New York: Holt, Rinehart & Winston, Inc., 1973.

McNeil, J. & Popham, J. The assessment of teacher competence. In Travers (Ed.), *Second handbook on research in teaching.* Chicago: Rand McNally & Co., 1973, 218-224.

Rosenshine, B. *Interpretive study of teaching behavior related to student achievement* (Final Report). College of Education, Temple University, 1970.

Rupley, W. Teacher instructional emphases and student achievement in reading. *Peabody Journal of Education,* in press.

Rutherford, W. Teacher improvement in reading. In Smith (Ed.), *Reading methods and teacher improvement.* Newark, Delaware: International Reading Association, 1971, 124-133.

Saadeh, J. Teacher effectiveness on classroom efficiency: A new teaching. *The Journal of Teacher Education,* 1970, *21,* 73-91.

ANN B. DOLGIN
University of Tampa

AN ANALYSIS OF BLACKS IN AMERICAN HISTORY TEXTBOOKS AS PERCEIVED BY SOCIAL STUDIES TEACHERS

Research indicates that curriculum materials may affect students' racial attitudes (Trager and Yarrow, 1952; Johnson, 1966). This study opens an area of research which examines content material so that further research in reading may be able to include the variable of racial attitude, which may be a factor affecting the learner's motivation to read a social studies textbook.

The purposes of the study were to develop an objective and reliable instrument for assessing the treatment of Blacks in secondary American history textbooks and to use that instrument to gather data on teacher perceptions of the treatment of Blacks in these books. A third purpose was to assess the adequacy of textbook content which deals with Black Americans.

STAGE 1: INSTRUMENT DEVELOPMENT

The investigator identified representative ideas about Black Americans based on the writings of historians and educators. These ideas were used to construct a content analysis instrument designed to test twenty-one textbooks used in the dual research study. The measuring instrument contained one hundred and four items in ten categories: slavery, abolitionism, discrimination, racial myths, contributions, desegregation, racial violence, racial harmony, events ending racial discrimination, and contemporary social and economic problems. The unit of analysis was the theme unit. A theme unit is an item of information in a single sentence which makes an assertion about the subject matter (Berelson, 1952).

A jury method was used to ascertain the validity of the instrument. Each expert worked independently. All eight judges accepted all categories as appropriate and judged 96 percent of the theme units to be appropriate.

STAGE 2: TEACHERS PERCEPTIONS OF BLACKS

The second purpose of this study was to measure teacher perceptions of information about Black Americans in textbooks. Specifically, the researcher attempted to determine whether those perceptions were mostly positive or negative, and whether they differed by coder variables such as race, sex, or years of teaching experience. Scores from the content analysis instrument were used as indicators of teacher perceptions of the treatment of Blacks in history texts.

Subjects

Sixty social studies teachers — thirty White and thirty Black — from Hillsborough County, Florida volunteered to serve as coders; i.e. to use the instrument to analyze an American history textbook. An orientation and training session were used to describe the purposes of the study. Specific hypotheses were kept from the coders on the assumption that such information might contaminate the findings. The training session focused on developing a standard form of reference for using the content analysis instrument.

Reliability

Kuder-Richardson formula 20 was used to obtain a reliability coefficient. This method was chosen because the content analysis instrument used in this study dealt with dichotomously scored items. An individual coder either agreed or disagreed that a theme unit was present in the text. For the fifteen texts, the obtained coefficients ranged from .70 to .96 with an average of .81.

Results and Discussion

Teachers rated textbooks as having significantly more positive than negative facts about Blacks ($x^2 = 15.78$, $p < .01$). Individual chi-squares were also calculated for each text. Three of the fifteen individual chi-squares were significant at the .01 level (book two: $x^2 = 7.54$, book eight: $x^2 = 16.96$, book eleven: $x^2 = 8.09$). In addition, the results for book thirteen were significant at the .05 level ($x^2 = 4.36$).

The investigator hypothesized that Black social studies teachers would identify more negative information about Blacks than would White teachers. Contrary to this expectation, White teachers identified significantly more negative content about Blacks than did Black teachers ($x^2 = 6.46, p < .05$).

Years of teaching experience of the teacher coders was not significantly related to the amount of negative versus positive textbook information which they identify using the content analysis instrument ($x^2 = 1.29, p > .05$).

Male and female teachers did not differ in the amount of perceived negative information about Black Americans in textbooks ($x^2 = 1.04, p > .05$).

Perhaps the most interesting finding pertains to the second hypothesis. Counter to expectation, White teachers identified more negative statements about Afro-Americans than Black teachers. These differences may be attributed to Blacks being more sensitive to the theme units themselves. Evidence of this was noted when the researcher spoke individually to each coder.

STAGE 3: ANALYSES OF TEXTS

The third objective of this study was to assess the adequacy of that portion of the content of textbooks which deals with Blacks. Six texts not used in the previous portion of the study were coded by the researcher with the instrument. A descriptive essay was written using the judgments of the investigator along with an evaluation of how these facts were integrated and historically treated within the content of each textbook. The findings were:

1. Textbook authors have increased historical references. A few books were excellent, but not one book provided comprehensive treatment of all major categories.
2. Although historical references about Blacks have increased, integration of these ideas is achieved by adding a new paragraph, a new page, or a new chapter with the exception of topics like slavery, abolitionism, and reconstruction.
3. Textbook authors generally portray Blacks with more dignity, but occasional inaccurate historical references still pervade a few books, inadvertently reinforcing stereotypes of Black people incapable of helping themselves.

Based upon previous research, textbooks are moving in the direction of improved presentation. This study revealed that revisions are still needed. For example, including the fact that Blacks actively fought for their freedom during the 1850's may help a student become sensitive to a need to read and reflect about clarifying his or her racial attitude, self-concept, and value orientation.

CONCLUSION

A review of the literature revealed that generally content analysis studies have not been used to measure perceptions of coders (Carpenter, 1941; American Council on Education, 1949; Marcus, 1961; Banks, 1969; and Kane, 1970). The findings of this study indicate that the variable of race is significant. Therefore, committee textbook evaluation teams should provide opportunities for White and Black educators to communicate their views openly so that personal perceptions may not interfere with objective assessment.

Copies of the instrument may be obtained from Ann B. Dolgin, University of Tampa, 401 West Kennedy Boulevard, Tampa, Florida 33606.

REFERENCES

Banks, J. A. *A content analysis of elementary American history textbooks: The treatment of the Negro and race relations.* Unpublished doctoral dissertation, Michigan State University, 1969.

Berelson, B. *Content analysis in communication research.* Glencoe, Illinois: The Free Press, 1952.

Carpenter, M. E. *The treatment of the Negro in American history school textbooks.* Menasha, Wisconsin: George Banta Publishing Co., 1941.

Committee on the Study of Teaching Materials in Intergroup Education. *Intergroup relations in teaching materials.* Washington, D.C.: American Council on Education, 1949.

Johnson, D. Freedom school effectiveness: Changes in attitudes of Negro children. *The Journal of Applied Behavioral Science.* 1966, *2*, 325-330.

Kane, M. B. *Minorities in textbooks: A study of their treatment in social studies texts.* Chicago: Quadrangle Books for Anti-Defamation League of B'nai B'rith, 1970.

Marcus, L. *The treatment of minorities in secondary school textbooks.* New York: Anti-Defamation League of B'nai B'rith, 1961.

Trager, H. & Yarrow, M. *They learn what they live.* New York: Harper & Brothers, 1952.

KATHLEEN STUMPF JONGSMA
Texas Woman's University, Denton, TX

READING TEACHERS' KNOWLEDGE OF AND ATTITUDES TOWARD BLACK ENGLISH

Several studies (Gilberts, Guckin, and Leeds, 1971; Naremore, 1970, 1971; Williams, Whitehead and Traupman, 1970a, 1970b; and Williams, Whitehead, and Miller, 1971, 1972) have found that teachers form opinions about speakers' ethnicities, backgrounds, personalities, and educational and occupational potentialities based on the hearing of very short samples of language behavior. While it is not always clear what is being perceived while listening, dialect, voice quality, or some other unknown, it is clear that speech characteristics evoke reactions in the listener so that the listener is able to respond definitively to written statements characterizing various personal and educational aspects of a speaker.

Other studies by various researchers (Labov, 1964, 1966; Morgan, 1969; Pederson, 1972; Sutton, 1970; Wiggins, 1970; and Williamson, 1968), have documented the phonological, morphological, and syntactic features found in Black English. Little, however, has been reported as to whether or not classroom teachers are knowledgeable about such features. Much classroom observation and teacher conversation suggests that teachers' understanding of dialect features is limited. At the time of the present study, no studies had been located which focused on the interactions among attitudes, knowledge, and behaviors.

The purpose of the present study, then, was to determine the relationships among teachers' knowledge of Black English features, teachers' attitudes toward the use of these dialect features, and teachers' verbal behavior toward dialect speakers in the classroom.

METHOD

Subjects

Thirty third- and fourth-grade female reading-language arts teachers served as subjects for the study. Subjects were selected from eight deliberately chosen elementary schools in the metropolitan area of New Orleans, Louisiana. Schools were selected to be representative of the racial, social, and organizational patterns of the metroplex. Subjects were selected for participation by building principals. Racial characteristics of the sample group, 80% White, 20% Black, matched the proportion of Whites to Blacks in the total school system.

Materials

Three separate instruments were constructed by the investigator. One assessed teachers' attitudes toward Black dialect, one assessed teachers' knowledge of

Black dialect, and a third assessed teachers' behavior toward Black dialect speakers when dialect speakers used their dialect in the classroom. The first two instruments required the use of language samples. Samples were obtained by a Black female bidialectal speaker, using stimuli developed by Shuy, Wolfram, and Riley (1967). Four final test tapes, each featuring the voice of one Black dialect speaker, were selected.

The first 1 1/2 minutes of each tape were used for assessing teacher language attitudes. A 25 item five-point Likert-type scale was developed. Items assessed attitudes toward the speaker's verbal ability, his teachability, his personal traits, and his family background. Items were derived from semantic differentials developed by Williams, et. al. (1970, 1971, 1972), and Naremore (1970, 1971); from comments in the literature about Black English speakers; and from conversations with teachers who presently teach Black English speakers.

The remaining three minutes of each tape segment were used to assess teachers' recognition of Black English features. All tapes were typescripted and features tabulated to serve as a master list against which teachers' responses were scored.

The third instrument, the Teachers' Language Behavior Observation Scale, was developed using guides suggested by Good and Brophy (1973), Madge (1965), and Mitzel (1963).

Reliability and validity were established by administering the instruments to 60 graduate and undergraduate students enrolled in summer school classes at the University of New Orleans. For the attitude assessment, estimates for the four separate scales, with Spearman-Brown corrections, ranged from .83 to .90. Concurrent validity with a similar attitude assessment scale used in a similar situation (Williams, et. al., 1970a) was assessed as .86.

For the scale assessing knowledge of features, reliability was determined according to an analysis of variance procedure suggested by Cureton (1951). Between-raters variance was included in the error term following Ebel's suggestions (1951) for inclusion when decisions are made by comparing single raw scores assigned to different pupils by different raters. The coefficient obtained, .41, was based on feature recognition of 60 subjects identifying Black English features for three of the four speakers.

The observation scale was submitted to colleagues involved in urban teacher preparation. All felt the instrument had sufficient face validity. Reliability of the instrument could not be determined by conventional means. By doing all necessary observation herself, the investigator eliminated the problem of inter-rater but not intra-rater reliability.

Procedures

All participants were observed for a minimum of three observations, averaging one hour each. While observing, the investigator coded each instance of teacher reaction to the language of Black English speakers, both verbal and nonverbal. No attempts, however, were made to record teachers' reactions to Standard English speakers as base line data.

Attitude and knowledge scales were completed after all observations had taken place. For the attitude scale, each tape was played for one- and a half-

88

minutes. After each segment was heard, the corresponding Likert scale was administered. After all attitude scales had been completed, teachers heard three one-minute speech segments for each of the four speakers. Teachers were told to write down as they listened any features which they considered characteristic of Black English.

RESULTS

Knowledge Scale

Overall performance indicated that the teachers recognized less than 12% of the total Black English features. Of the total features recognized, *th* substitutions (*dat* for *that*, *den* for *then*), especially in the initial position, were most frequently recognized. Also easily identified were subject-verb differences, especially in third person singular, present tense (she *come* back). Other features recognized with some frequency included vowel substitutions (*pernt* for *point*, *foist* for *first*), vowel lengthening (*daid* for *dead*), and simplification of final consonant clusters (*sen'* for *send*, *las'* for *last*). Overall, teachers identified more morphological features than they did phonological features.

Attitude Scale

Attitude scores could range from 25 to 125 for each speaker, with pilot data indicating that 75 was the breakpoint for negative or positive scores. Differences in item means and standard deviations (Speaker One, $M = 86.97$, $SD = 12.10$; Speaker Two, $M = 65.13$, $SD = 10.65$; Speaker Three, $M = 77.30$, $SD = 13.40$; and Speaker Four, $M = 88.33$, $SD = 11.21$), suggest that teachers rated speakers differently, varying their scores with the different stimuli heard. Mildly positive attitudes were noted for three of the four speakers.

Observation Scale

While teachers made more positive comments than they made negative or neutral, mean responses in any category were few (Total positives, $M = 15.63$, $SD = 10.80$; Total negatives, $M = 1.03$, $SD = 2.59$; Total Neutrals, $M = 1.33$, $SD = 1.83$), indicating that few responses, either verbal or nonverbal were made to Black English speakers during the observation.

Relationships Among the Criterion Variables

It was found that there was no significant relationship ($r = .14$) between teachers' attitudes toward Black English and their positive comments to Black English speakers. Neither was there a significant relationship ($r = -.10$) between teachers' knowledge of Black English and their comments to Black English speakers. There was, however, a significant relationship ($r = -.55$, $p < .01$) between knowledge of Black English and teachers' attitudes toward Black English speakers.

After visual inspection of the scatter diagrams suggested a nonlinear relationship, supplementary data analysis procedures were followed. Subsequent testing of the significance of the eta coefficient (eta $= .77$, $F(9, 20) = 3.271$, $p < .025$), the linear correlation coefficient ($r = -.11$, $F(1, 28) = .311$, *n. s.*), and the nonlinearity of regression ($F(8, 20) = 3.578$, $p < .01$) demonstrated that the linear coefficients previously obtained were not suitable for describing the relationship between knowledge and attitudes and that the eta coefficient better described the degree of association between teachers' scores on the two variables. In the present study, the degree of association between knowledge and attitude varies. Those teachers who had least knowledge of Black English had neutral or slightly positive attitudes toward Black English features. Teachers with slightly more knowledge had more positive attitudes. After a point, however, increases in knowledge were not associated with increases in attitudes, and attitudes declined. Any attempt to interpret the specific form of the curvilinear relationship, however, must be treated cautiously because of the cellar effect of the knowledge scores, for most knowledge scores were clustered at the low end of the scale.

Supplementary Data Analysis

Mann Whitney U tests were calculated among the teachers' scores on the three criterion variables when considering the variates of race, experience, and education.

Significant differences ($z = -1.9847$, $p < .024$) were found between Black and White teachers on the attitude scales. Black teachers ($M = 342.5$) had more positive attitudes toward dialect speakers than did White teachers ($M = 311.5$). Significant differences ($z = 1.8036$, $p < .036$) were found between Black teachers and White teachers on the behavior scale. White teachers ($M = 16.83$) had more positive classroom responses toward Black English speakers than did Black teachers ($M = 10.83$). No significant differences were found, however, on the knowledge scale.

Considering the variate of level of higher education completed, no significant differences were found between those teachers holding bachelors' degrees and those holding masters' degrees. Neither were significant differences found in teacher behavior or teacher attitude according to years of teacher experience. The H (9.87) calculated for teachers' knowledge was, however, significant ($p < .02$). There were significant differences in knowledge according to years of experience. Those teachers with experience of 6-10 years had more knowledge about the dialect ($M = 48.60$) than did those with 1-5 years ($M = 30.79$), 11-15 years ($M = 32.25$) or more than 16 years ($M = 25$).

DISCUSSION

The present study confirms findings of Naremore (1970, 1971), and Bronstein, et. al. (1970), in that it again demonstrates that Black teachers and White teachers frequently differ in the favorableness of their ratings of Black English speakers. Black teachers seemed to have more positive attitudes toward speakers of Black English than do White.

The present study also suggests that teachers are unable to identify many of the features linguists cite as representative of Black English. As other studies had not been done in the area of teachers' feature identification, the present percentages and patterns of identification suggest that teachers have considerable difficulty identifying common phonological and morphological features. An added dimension concerning knowledge of features also emerged in the present study. Williams, et. al. (1972) had previously found no differences with attitudes according to experience, and this finding was confirmed in the present study; but the differences in knowledge introduce a new consideration as to when teaching about dialect might be most meaningful to teachers.

If, after replication, similar findings were obtained, the investigator would recommend that teachers be introduced to programs about Black dialect features. The present study indicates that teachers are more open to Black English speakers when they know something about phonological, morphological, and syntactic features of Black English.

REFERENCES

Bronstein, A. J., Dubner, F., Lee, P., & Raphael, L. J. *A sociolinguistic comment on the changing attitudes toward the use of Black English and an experimental study to measure some of these attitudes.* Paper presented at the annual convention of the Speech Communication Association, New Orleans, December 1970. (ERIC Document Reproduction Service No. ED 051 226).

Cureton, E. E. *Reliability of a set of averages of unequal numbers of scores.* Paper presented at the meeting of the Southern Society for Philosophy and Psychology, March, 1951.

Ebel, R. L. Estimation of the reliability of ratings. *Psychometrika,* 1951, *16,* 407-424.

Gilberts, R. A., Guckin, J. P., & Leeds, D. S. *Teacher perceptions of race, socio-economic status and language characteristics.* Paper presented at the annual meeting of the American Educational Research Association, New York, February, 1971. (ERIC Document Reproduction Service No. ED 052 131).

Good, T. L. & Brophy, J. E. *Looking in classrooms.* New York: Harper and Row, Publishers, 1973.

Labov, W. Stages in the acquisition of Standard English. In R. W. Shuy, (Ed.), *Social dialects and language learning.* Champaign: National Council of Teachers of English, 1964, 77-104.

Labov, W. *The social stratification of English in New York City.* Washington, D. C.: Center for Applied Linguistics, 1966.

Madge, J. *The tools of social science.* Garden City, New York: Anchor Books, Doubleday and Company, Inc., 1965.

Medley, D. M. & Mitzel, H. E. Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), *Handbook of research on teaching.* Chicago: Rand McNally, 1963, 247-329.

Morgan, A. L. *An analytical survey of the oral language patterns of culturally disadvantaged children of the New Orleans area.* Unpublished masters thesis, Louisiana State University in New Orleans, 1969.

Naremore, R. C. *Teacher differences in attitudes toward children's speech characteristics.* Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, March 6, 1970. (ERIC Document Reproduction Service No. ED 039 307).

Naremore, R. C. Teachers' judgments of children's speech: A factor analytic study of attitudes. *Speech Monographs,* 1971, *38,* 17-27.

Pederson, L. A. The pronunciation of English in Metropolitan Chicago. In J. Williamson and V. Burke (Eds.), *A various language: Perspectives on a theme.* New York: Holt, Rinehart, and Winston, 1972, 525-548.

Shuy, R., Wolfram, W., & Riley, W. *Linguistic correlates of social stratification in Detroit speech.* Michigan State University, 1967.

Sutton, L. H. *An analysis of oral language responses of culturally disadvantaged Negro children of first grade level in the New Orleans area.* Unpublished masters thesis, Louisiana State University in New Orleans, 1970.

Wiggins, A. *A study of dialect differences in the speech of first grade Negro children in the inner city schools of Cleveland, Ohio.* Unpublished doctoral dissertation, Indiana University, 1970.

Williams, F., Whitehead, J. L., & Traupman, J. *Correspondence between semantic differential ratings of children's speech and speech anticipated upon the basis of stereotype.* Technical Report, Center for Communication Research, 1970a. (ERIC Document Reproduction Service No. ED 042 758).

Williams, F., Whitehead, J. L., & Traupman, J. *Latency of teachers' semantic differential ratings of children's speech.* Center for Communication Research, 1970b. (ERIC Document Reproduction Service No. ED 042 757).

Williams, F., Whitehead, J. L., & Miller, L. M. Ethnic stereotyping and judgments of children's speech. *Speech Monographs,* 1971, *38,* 166-171.

Williams, F., Whitehead, J. L., & Miller, L. M. Relations between language attitudes and teacher expectancy. *American Educational Research Journal,* 1972, *9,* 263-279.

Williamson, J. W. A phonological and morphological study of the speech of the Negro of Memphis, Tennessee. *Publication of the American Dialect Society,* 1968, *50,* 32-45.

HAL W. SEATON
University of Georgia

SHARON D. SMITH
M-CESA, Atlanta, GA

THE USE OF RECEPTIVE AND EXPRESSIVE LANGUAGE DEVELOPMENT ACTIVITIES FOR IMPROVING STUDENTS' READING ACHIEVEMENT

The relationships among various language activities often are not understood by teachers who continually attempt to influence different aspects of children's language. Teachers, based on observations of their students, assume that increased facility with one facet of language enhances other language uses.

The capacity for language acquisition is often considered universal. Lenneberg (1973) argued that language is innate behavior for human beings. Generally, language is expected among people; any deficiency is considered abnormal.

The problems teachers face in coordinating oral language and reading instruction result from a lack of instruction in teacher preparation programs, according to Moore (1969). Another factor of the problem is the fact that students enter school with speaking and listening skills, so that teachers are not required to study or to encourage the development of these language skills (Loban, 1969).

Loban proceeded to report confirmation from teachers' observations and from research that language abilities are interrelated. Teachers have discovered that outstanding readers and writers also excel in spoken language production (1969, pp. 101-102).

Currently, teachers are encouraged to create classroom situations which simulate the natural environment in which oral language develops, as a means of developing new language skills or improving existing ones (Cazden, 1972; Tovey, 1976).

This study was designed to improve children's receptive and expressive language. Its further purpose was to investigate the effects of a two-year systematic language development program on students' reading ability.

METHOD

Subjects

One hundred ninety-two students in a public school kindergarten were randomly assigned to either the experimental or the control group. The sample included 133 White and 59 Black students, primarily on the lower economic levels. The experimental group included 44 males and 49 females; the control group 47 males and 52 females.

Pretests indicated the groups were not significantly different in terms of chronological age, intelligence, auditory comprehension, and verbal ability. Group means from the *Cognitive Abilities Test* (Thorndike, Hagen, & Lorge, 1965) and the *Preschool Language Scale* (Zimmerman, Steiner, & Evatt, 1969) are presented in Table 1.

Table 1

Mean Scores by Groups on Pretests of Intelligence and Language

	IQ				Language in Months		
Group	*N*	*M*	*SD*	*t*	*M*	*SD*	*t*
Experimental	93	94.49	11.96	0.12	55.30	9.81	-0.09
Control	99	94.29	10.44		55.42	9.06	

Procedure

The experimental group received eighty hours of directed instruction during four six-week periods, while the control group spent equivalent time in unstructured language activities. The directed instruction was in areas of (1) verbal expression of ideas and feelings, (2) visual perception and memory, (3) auditory perception and memory, and (4) conceptualization. In each area, beginning tasks were simple; complexity increased over the six week period. The instruction took place in forty-minute daily segments, of six months' duration, with single activities extending ten to fifteen minutes. On occasion, slightly longer time was allowed, if students were highly motivated and the activity was holding students' attention. Game formats and group experiences generated many of the learning and practice situations.

Verbal expression of ideas and feeling. The object of developing children's verbal expression was to increase the occurrence of verbal exchange between the students and others, as well as to enhance the type of verbal interaction. Initial activities involved each student reciting name, age, and family members. Gradually new concepts such as size, color, location were introduced as pupils described pets, self, home, and school, among other objects. Guessing games such as "I Spy" and "I am thinking of . . ." were used. After describing things, students later reconstructed events, repeated jokes, and explained television programs. The next step was to have students create original situations to describe, beginning with pictures as stimuli. Teachers expanded the words and concepts used by adding and explaining new items. Eventually students practiced sequencing events, using complete sentences, and inflecting word forms.

Visual perception and memory. The second area of instruction directed student's practice to visual perception and memory. The experimental group worked with geometric designs, naming, describing and drawing those designs using templates. Some of the concepts introduced during the initial period of instruction were further developed. In game-playing situations, students looked at groups of objects, then identified what was removed and gave descriptions for the missing objects. Known objects were described in terms of geometric configuration, in addition to the concepts of size and color previously taught.

Auditory perception and memory. The third segment of the program again focused on training perception and memory, but in the auditory realm many recordings, often teacher-made, were used to allow children to guess the source of the sounds. At times, descriptive hints were offered. Another skill involved was following oral directions, moving from one command to two, then three. Lessons in auditory discrimination progressed from gross to fine differences, culminating with the use of individual letter sounds. Frequently used exercises including rhyming words, recalling sequence, retelling stories, discussion of stories, and predicting or identifying possible outcomes.

Conceptualization. The final portion of the language development program dealt with conceptualization. The experimental group received training in identifying opposites using 75 word pairs; classifying objects by color, number, shape, or group; and counting up to 10 concrete objects. Other practice involved associated words (cat and kitten, salt and pepper, hammer and nails), common prepositions, homonyms, and words with multiple meanings. Word analogies, comparative language structures, and abstract ideas were included in the activities.

Testing Instruments

At the end of one year in kindergarten, subjects were tested using the *Preschool Language Scale*, the *Metropolitan Readiness Test* (Hildreth, Griffiths, & McGauvran, 1969), and the *Cognitive Abilities Test*. Further evaluation at the end of first grade was accomplished by means of the *Metropolitan Achievement Tests: Reading Tests* (Durost, et al., 1971).

Statistical Analysis

The data were analyzed using a two-tailed *t*-test to determine the significance of the differences between group means on pre- and post-measures of intelligence and language, including auditory comprehension and verbal ability. Other differences compared were chronological age, reading readiness, word knowledge, word analysis techniques and reading comprehension.

RESULTS

After one year, a significant difference favoring the experimental group was found on scores obtained from the *Preschool Language Scale* (scores indicated in Table 2 combine both subtests) and the *Metropolitan Readiness Test*. No significant difference in chronological age or in scores from the post-test administration of the *Cognitive Abilities Test* existed between groups. See Table 2.

Table 2

Mean Scores by Groups for 1st Year Tests

		IQ			Language in Months			Readiness in Months		
Group	*N*	*M*	*SD*	*t*	*M*	*SD*	*t*	*M*	*SD*	*t*
Experimental	93	103.15	12.20	1.7	69.95	9.46	4.19*	64.59	12.37	7.24*
Control	99	100.32	10.86		64.01	10.20		50.49	14.58	

*$p < .01$

After two years (at the end of first grade), students in the experimental group performed significantly better on the *Metropolitan Achievement Tests: Reading Tests*, as shown in Table 3.

Table 3

Mean Scores by Groups for 2nd Year Tests

		Word Knowledge			Word Analysis			Reading Comprehension		
Group	*N*	*M*	*SD*	*t*	*M*	*SD*	*t*	*M*	*SD*	*t*
Experimental	83	44.01	12.90	2.77*	39.08	10.33	4.94*	43.92	13.00	3.24*
Control	87	38.82	11.44		31.95	8.32		38.16	9.84	

*$p < .01$

Comparisons by sex made between groups indicated that experimental males were significantly better than control males on the posttest language measure, $t(89) = 4.46, p < .01$, and, similarly, experimental females were better, $t(99) = 3.00, p < .01$.

Further analysis (see Table 3 for *t* values) showed that within the experimental group, females yielded significantly higher scores than males on measures of language, readiness, word knowledge, word analysis, and reading comprehension. Also, when students within the experimental group were sub-grouped according to IQ, (70 - 100, above 100), subjects in the higher level of intelligence made significantly higher scores on posttest measures of language, readiness, word knowledge, word analysis, and reading comprehension. Another difference within the experimental group was that students older at first grade entry made significantly higher scores on both the pre- and post measures of the *Preschool Language Scale*, $t(91) = -2.78, p < .01, t(91) = -1.52, p < .05$.

Table 4

Posttest Results for Experimental Group[*]

Posttest	Sex Differences[+]	IQ Differences
Language	$t\,(91) = -2.88$	$t\,(83) = -5.43$
Readiness	$t\,(91) = -4.83$	$t\,(83) = -8.45$
Word Knowledge	$t\,(81) = -5.52$	$t\,(73) = -6.89$
Word Analysis	$t\,(81) = -5.49$	$t\,(73) = -6.43$
Reading Comprehension	$t\,(81) = -5.42$	$t\,(73) = -6.76$

[*] $p < .01$ in all cases

[+] in each case, females outperformed males

CONCLUSIONS

The results of this study indicate that including a systematic program of language development in early instructional programs should increase reading achievement. The findings affirm Cazden's (1972) recommendation that children need a school environment that allows many opportunities for communication. One of the implications of these findings is that, while students fare well in informal sessions of language interaction, teachers must be aware of and consciously direct activities toward developing specific language-related abilities. Associated with that implication is the finding that reading readiness can be positively affected by planned programs of language development, especially among children who come to school less well prepared by their backgrounds for reading. Furthermore, this study lends suppport to the belief that reading and language are truly related.

REFERENCES

Cazden, C. B. *Child language and education.* New York: Holt, Rinehart and Winston, 1972.

Durost, W. N., Bixler, H. H., Wrightstone, J. W., Prescott, G. A., & Balow, I. H. *Metropolitan Achievement Tests: Reading Tests.* New York: Harcourt Brace Jovanovich, 1971.

Hildreth, G. H., Grifiths, N. L., & McGauvran, M. E. *Metropolitan Readiness Tests.* New York: Harcourt Brace Jovanovich, 1969.

Lenneberg, E. H. The capacity for language acquisition. In M. Lester (Ed.), *Readings in applied transformational grammar.* New York: Holt, Rinehart, and Winston, 1973.

Loban, W. Oral language and learning. In J. Walden (Ed.), *Oral language and reading.* Champaign, Ill.: National Council of Teachers of English, 1969.

Miller, W. R. Language acquisition and reading. In J. Walden (Ed.), *Oral language and reading.* Champaign, Ill.: National Council of Teachers of English, 1969.

Moore, W. J. Oral language and reading: an introductory comment. In J. Walden (Ed.), *Oral language and reading.* Champaign, Ill.: National Council of Teachers of English, 1969.

Thorndike, R. L., Hagen, E., & Lorge, I. *Cognitive Abilities Test.* New York: Houghton, Mifflin, 1968.

Tovey, D. R. Language acquisition: A key to effective language instruction. *Elementary English,* 1976, *53,* 868-873.

Zimmerman, I. L., Steiner, V. G., & Evatt, R. L. *Preschool Language Scale.* Columbus, Ohio: Charles E. Merrill, 1969.

ANABEL P. NEWMAN
Indiana University

TWELVE YEAR STUDY OF PUPILS WHO WERE UNDERACHIEVING IN READING FIRST GRADE

The Cedar Rapids' study (Reid and Beltramo, 1965) was one of the original First Grade Studies (Bond and Dykstra, 1967a and b). Its uniqueness lay in the fact that it sought solutions to the reading problems of the so-called low-group pupil. Extensive lesson plans particularly suited to the needs of the lower third of the first grade pupils in Cedar Rapids, Iowa were developed. Many of the recommendations appearing in the original Cedar Rapids' final report have since been incorporated into published materials.

Significant also from the Cedar Rapids' study were the reactions of both teachers and pupils. Teachers said that they had not realized that low-group pupils could be so overwhelmingly responsive. Teachers recommended strongly providing *different* materials (not just the same materials at a slower pace) for the low-group youngster. The children had had attractive materials with which they were consistently successful, and consequently had not experienced the disheartening lack of regard which many low-group first graders meet when they come to school expecting to learn to read.

A sixth grade followup of pupils participating in the Cedar Rapids' first grade study was conducted by Newman in which it was found that ". . . random assignment of treatment to classroom has neither short nor long-range effect on later learning characteristics of low-group youngsters (1971, p. 60)." It was also found that first grade achievement far outweighed first grade readiness as a predictive measure, and that a strong *g*-factor existed in first and sixth grade reading achievement. Findings concerning sex differences in learning were noted. Most significantly, at least one-third of the children originally designated as low-group readiness were achieving at or above grade level in reading in sixth grade.

A ninth grade followup study of pupils who had been part of the Reid and Beltramo first grade sample was conducted utilizing scores on the Iowa Test of Educational Development (ITED) as criterion. Results are available in two unpublished studies (Newman and Lohnes, 1974, 1975), and indicate (a) that the diagnosis of low readiness upon entry to first grade does suggest the likelihood of permanent deficit in general intellectual development, even if the child receives competent teaching during the primary school years; and (b) that the degree of severity of the deficit does forecast the level of general intellectual development assessed at the close of grade nine.

The study reported herein presents preliminary results of the eleventh and twelfth grade portion of a longitudinal investigation of Cedar Rapids, Iowa pupils who were low readiness in reading in first grade (1964-65). Questions pursued in this study were as follows: (a) What are the educational, and insofar as possible, the psychological, physiological, and cultural characteristics of a selected group of high school juniors and seniors who were low readiness in first grade? and (b) Which variables are associated with the achievement of such low readiness pupils over a twelve year period? In terms of design, the characteristics of Achievers and Low Achievers interviewed from among a group of students who were *all* originally low readiness in first grade were contrasted.

The ninth grade Iowa Tests of Educational Development (ITED) scores of all students who participated in the original first grade study were examined ($N = 135$). All students scoring at or above cohort mean were, for the purposes of this study, considered to be Achievers ($N = 89$). Those who did not score at or above cohort mean on any subtest of the ITED were classified as Low Achievers ($N = 46$).

Student interviews (taped in response to an 18 item questionnaire and telephone interviews of parents (12 item questionnaire) were personally conducted by the investigator for one week periods in May, 1975 and January, 1976. It was possible to complete 71 student interviews and 72 parent interviews during these two, one-week periods.[1]

Fifty-two of the interviews were from the pool of 89 students (58%) who achieved at or above grade level on at least one subtest of the ITED in ninth grade (Achievers). Nineteen interviews were from the pool of 46 students (41%) who did not achieve at or above cohort mean on a single subtest of the ITED (Low Achievers).

RESULTS

Responses from several questions in the student and parent interviews provide a heuristic look at the data in terms of answering question one above regarding educational, psychological, physiological, and cultural considerations. Psychologically, Achievers scored consistently higher on IQ tests at 1st, 6th, and 9th grades. Physiologically, Low Achievers were beset by more physical problems of a magnitude likely to affect school achievement than were Achievers. Culturally, mean income was higher for Achieving students. Parent educational level was also higher for Achieving students.

In response to question two, regarding variables associated with achievement of low readiness pupils over a twelve year period, the following trends have emerged to date: Student memories of grade school and high school were positive for both groups; the subject area preferred by both groups was Reading/Language Arts. Interests are wide in both groups, and many will become career commitments. Sixty-eight percent of Achievers interviewed enjoy

---

[1] Although the numbers of students and parents interviewed seem similar, they do not necessarily reflect relationship. Student interviews were conducted during school hours. Parent interviews were conducted after school, and from five to ten p.m. each evening. It was not always possible, however, to reach parents of students interviewed. Interviews were obtained with parents of any of the 135 students who could be reached.

100

reading. Low Achievers interviewed enjoy reading if it's something which interests them.

Both groups of students are from generally stable homes with nuclear familes: father, mother, and siblings. Ninety-three percent of Achievers perceived home influences operating on their school achievement. With Low Achievers, the response was about 50-50. Peer influences were more generally recognized and credited as positively influencing for Achievers. Seventy-two percent of Achievers see themselves as college bound. Low Achievers see vocational training as their career goal after high school. In both groups, there seems to be a variety of variables (family, friends, a special teacher, the opportunity to work or travel) contributing to the gains of the students.

Comparison of the student responses on five specific variables: model, motivation, interest, perseverance, and pressure reveals a preponderance of emphasis by parents of Achieving students on model and motivation, with little emphasis by these parents on pressure. In contrast, fifty percent of the parents of Low Achieving students mentioned some element of pressure being exerted to get their children to achieve. These same parents of Low Achieving students did not seem to emphasize model or motivation as much as did the parents of the Achieving students.

Low achieving student responses regarding motivation fell under the headings of pressure or influence. By contrast, responses of the Achieving students fell into categories of praise, pay, and competition. Thus, although fifty percent of the Achieving responses were externally motivated, they perceived their motivation to be somewhat different from that of the Low Achievers.

In terms of internal motivation, the Achieving students made a strong showing on an item indicative of inner drive, while the Low Achieving students indicated that interest and desire tended to be the more important motivator. Perseverance seemed more understood by and identified with Achievers.

Both Achievers and Low Achievers indicated strong interests in the category of sports/hobbies. Achievers, however, reported proportionably larger numbers of hobbies and art/crafts interests than did Low Achievers.

CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

If one perceives teaching/learning as a decision-making process on the part of teachers, then it would appear from the findings above that at least two major decision areas, strategies and resources, must be carefully considered in the instruction of youngsters who may seem to be low readiness in the primary grades.

Strategies

Model and motivation were important variables in the educational development and achievement of these Achieving low readiness students. It would seem important to develop strategies which incorporate significant amounts of modeling and motivation for such students. The role of extrinsic versus intrinsic motivation with externally and internally motivated students should also be considered, and finally, strategies for demonstrating the importance of perseverance to low achievement youngsters should also be explored.

Resources

In viewing the data provided on these youngsters over a 12 year period, it would appear that almost all of those interviewed are pursuing a course which will be productive to themselves and society. They are, by virtue of their arrival at eleventh and twelfth grade in Cedar Rapids schools, readers, whether they scored positively on standardized tests or not. Their interests are well developed and even though they might not all characterize themselves as avid readers, even the students in the low achieving group recognize and pursue reading as a skill leading to desirable ends.

These results would suggest the possibility that even greater use should be made of student interests in exploring resources that would stimulate students to further growth in reading.

Currently a study is underway, sponsored by a Spencer Foundation Grant, which will present case studies of 20 students whose eleventh and twelfth grade performance showed "the greatest contrast with original first grade prognosis." This study will highlight those aspects of model, motivation, interest perseverance, and pressure which characterized each student and will present in detail the psychological, physiological, cultural, and educational characteristics of each student profiled. It will be possible at that time to compare and contrast the achievement status of this smaller group of pupils with the 71 students whose data have been presented in the preceding preliminary study.

REFERENCES

Bond, G. L. & Dykstra, R. The cooperative research program in first-grade reading instruction. *Reading Research Quarterly*, 1967a, *11*, 5-141.

Bond, G. L., & Dykstra, R. *Final report Project X-001, Contract OE-5-10-264*, 1967b.

Eller, W. Contribution of the first and second grade studies. In J. A. Figural (Ed.), *Reading and realism. 13* (part 1) Proceedings of the 13th annual convention. Newark, Delaware: International Reading Convention, 1969.

Lindquist, E. F. and Feldt, L. S., *Iowa Tests of Educational Development* Chicago: Science Research Associates, 1970.

Newman, A. *Longitudinal study of pupils who were underachieving in reading.* Unpublished Ed.D. dissertation, State University of New York at Buffalo, 1971.

Newman, A. Later achievement of pupils underachieving in reading in first grade. *Reading Research Quarterly*, 1972, *7*, 477-50

Newman, A. & Lohnes, P., *Developmental trends in low reading readiness pupils over nine years.* Unpublished manuscript, 1974.

Newman, A. & Lohnes, P., *Correlations of G-type factors for low readiness pupils over nine years.* Unpublished manuscript, 1975.

Reid, H. C. & Beltramo, L., *The effect of different approaches of initial instruction on the reading achievement of a selected group of first grade children.* USOE Cooperative Research Project No. 2698. Cedar Rapids, Iowa: Cedar Rapids Community School District, 1965.

EARL F. RANKIN
University of Kentucky

SEQUENCE STRATEGIES FOR TEACHING READING COMPREHENSION WITH THE CLOZE PROCEDURE

The purpose of this paper is to present some concrete, practical suggestions for using the cloze procedure more effectively in teaching reading comprehension. Rankin (1957) was one of the first to call for the use of the cloze procedure as an instructional tool in reading. Unfortunately, although a number of attempts have been made over the years to improve reading comprehension through the use of the cloze procedure, results have been generally disappointing (Jongsma, 1971). In the light of both research evidence and theory, we have every reason to expect more positive outcomes. In a survey of factor analyses carried out on reading test results, Rankin (1962) concluded that two factors account for most of the variance in measures of reading comprehension: (a) vocabulary and (b) reasoning in reading. A rational analysis of cognitive processes involved in making cloze responses indicates that both of these two factors are operating. In addition, numerous high correlations between the results of cloze tests and standardized reading tests suggest that similar reading processes are being measured by both types of instruments.

A psycholinguistic theoretical view (Smith, 1971) of the reading process supports the expectation that the "cloze" might be useful in teaching reading. As Smith so cogently points out, two kinds of redundancy are involved in grasping the meaning of the printed passage: (a) distributional and (b) sequential. Redundancy is defined as information coming from more than one source. Distributional word redundancy is provided by the reader's general cognitive grasp of differential probabilities that certain words will occur in the English language independent of sentence context. For example, "cat" as a word is more likely to occur than "cloze" in English sentences. Sequential word redundancy is available to the reader who "expects" that certain words are more likely to occur than other words in a given verbal sequence. In the common expression, "No news is ____________ news," the missing word is more likely to be "good" than "bad." According to Smith, the fluent reader is able to utilize such redundant information stored in his long-term memory in addition to other types of information at a speed which is beyond the known limitations of nerves to transmit impulses. Since even a cursory examination of a cloze exercise would reveal that making correct cloze responses involves sensitivity to clues for both distributional and sequential redundancy, the use of cloze materials for training

---

The writer wishes to express his sincere appreciation to his research assistant, Ms. Linda Evers, for her invaluable editorial assistance in preparing this paper.

purposes in the reading classroom would appear to be a promising approach for developing this sensitivity.

However, many attempts to teach reading with the cloze procedure have consisted of a rather mechanical use of exercise materials with the hope that merely filling in blank spaces would somehow bring about better reading comprehension (Jongsma, 1971). It is the author's view that materials, per se, do not teach. Only teachers teach. Therefore, the cloze procedure can only be useful to the extent that it is used by a competent teacher in accordance with basic learning principles. The proper sequencing of instruction is fundamental to good teaching. Although little direct empirical evidence exists concerning the best sequence to use in teaching reading, it is nevertheless possible to suggest several instructional "sequence strategies" for using "cloze" in ways which might produce better results than have been obtained in most research. The term "sequence strategies" is used to denote teaching plans which may be useful in attaining an educational goal by organizing instruction in a given order of succession. The discussion of sequence strategies will be organized according to the following topics: I. Introducing the Cloze Procedure; II. Selecting Reading Passages; III. Choosing Scoring Procedures; IV. Selecting Word Deletions; V. Determining Response Types; VI. Using Visual Clues; and VII. Using Reinforcements.

These strategies should be considered as hypotheses to be tested rather than firm conclusions. They are stated as general propositions. Only empirical studies can determine whether or not they are applicable to different age levels or other designated groups.

The remaining discussion is restricted to a consideration of these sequence strategies. No attempt is made to discuss other important aspects of teaching with cloze such as use of discussion, various forms of feedback, etc.

I. INTRODUCING THE CLOZE PROCEDURE

The lack of "face validity" (i.e., acceptability by students) is an important deficiency of cloze materials in general. Such exercises are novel and threatening to many readers. In order to reduce the student's anxiety in working through a cloze exercise, the following instructional sequence is proposed:

SEQUENCE STRATEGY 1A

1) Use aural cloze.
2) Use aural-visual cloze.
3) Use visual cloze.

In an aural cloze exercise, the teacher reads orally a brief cloze passage, stopping at choice points where words are omitted, while the students write the deleted words in numbered order on a sheet of paper. It is advisable to provide a "beep" or bell sound to mark each deletion and then to proceed beyond that point to the end of the phrase or sentence before the written response is made by the student. Thus, bilateral constraints can operate to provide clues on both "sides" of the missing word. Sufficient time must be allowed by the teacher for the written responses to be made by even the slowest students. This technique should be used only to introduce the general nature of the cloze experience and should be accompanied by obvious vocal modulations and facial expressions. This will help compensate for the fact that the listener, unlike the reader, does not have an immediate chance to take another look at the clues surrounding the blank space. Of course, the teacher should repeat phrases or sentences on request to increase the probability of making correct responses.

An aural-visual cloze exercise is administered in the same way as previously described for administering the aural-cloze exercise except that, in addition, the students read the printed cloze passage silently while the teacher reads it orally. Thus, there are two sources of clues available to the student: (1) natural facial and vocal clues provided by the teacher, (2) the opportunity to read and re-read the context on either side of each missing word as in a conventional visual cloze exercise.

After practice with the previous two types of cloze presentations, most readers should be ready to work on conventional visual cloze materials with greater ease and less threat of novelty. In order to determine the amount of time needed by individuals on tasks at steps 1 and 2, the teacher must be sensitive to student performance at all times.

Sequence strategy IA is predicated on the assumption that filling in "gaps" in an oral communication is an experience which most people are more explicitly aware of than filling in visual "gaps" in a mutilated printed message. As such, this sequence is proposed as an introduction to a novel experience. Unless suitable pacing is used together with careful consideration of such variables as passage length, passage content, and provisions for an appropriate amount of context surrounding each pause, this procedure may overload the short-term memory of students, particularly young students.

II. SELECTING READING PASSAGES

The following sequence strategies employ common principles of proceeding from the known to the unknown, the easy to the difficult, and from the more interesting to the less interesting:

SEQUENCE STRATEGY IIA

1) Use language experience materials.
2) Proceed to materials from other sources.

SEQUENCE STRATEGY IIB

1) Use easy materials.
2) Proceed gradually to more difficult readability levels.

SEQUENCE STRATEGY IIC

1) Use high-interest, narrative materials.
2) Proceed to possibly less interesting, expository materials.

Space limitations will permit only a few comments on category II sequence strategies. In applying strategy IIA, the teacher will find language experience materials to be most useful in remedial work. The use of such materials for the production of cloze exercises has potential advantages not yet realized. With regard to strategy IIB, the transition from easy materials to more difficult materials can be facilitated by collecting many short passages and categorizing them by readability formula results. The application of strategy IIC is particularly appropriate for improving student motivation.

Cloze scores can be used to monitor transition from (1) language experience to other types of passages, (2) easy to more difficult passages, (3) more interesting to less interesting passages. The cloze procedure allows for great precision in monitoring the speed with which the student is able to make these transitions.

III. CHOOSING SCORING PROCEDURES

The acceptability of cloze procedures is greatly reduced when the teacher only gives credit for written responses which conform to the precise word deleted from the passage. Of necessity, this type of scoring produces a large number of mistakes. On the other hand, for superior readers with advanced knowledge of the subject matter in the passage or familiarity with the style of the author, this scoring procedure is challenging and provides a sensitive indication of small differences in achievement. For the teacher who wishes to use both types of scoring, the following sequence is recommended:

SEQUENCE STRATEGY IIIA

1) Use "synonym" scoring.
2) Use "exact word" scoring, if desired.

For instructional purposes, the determination of synonymity can be rather loosely made by the teacher and/or the students by asking the question, "Does this word make sense in the total context of the sentence and the surrounding paragraph?" In general, this is the best type of scoring procedure to use in teaching with the cloze procedure.

IV. SELECTING WORD DELETIONS

The following sequence strategies are concerned with the number and type of words to be deleted in the cloze exercise:

SEQUENCE STRATEGY IVA

1) Use a low deletion ratio (e.g., 1 deletion per 10 words).
2) Gradually proceed toward higher deletion ratios (e.g., 1 deletion per 5 words) using cloze scores to determine pace.

Regarding the use of strategy IVA, the lower the deletion ratio (i.e., number of deletions in relation to the number of running words), the more context will surround a given deletion. Greater context provides more redundancy clues for making responses. Thus, a reader may start with exercises constructed by deleting every tenth word and proceed toward higher ratios of every eighth word, every sixth word, etc. As he achieves some satisfactory criterion score at one deletion ratio, the reader can be given cloze materials of equivalent readability constructed at higher deletion ratios, which provide less context and therefore are more difficult. The ratio should not fall above 1 deletion per 5 words in order to maintain sufficient independence among cloze responses.

SEQUENCE STRATEGY IVB

1) Use "lexical" deletions with prompts.
2) Use "structural" deletions with prompts.

3) Use "any word" deletions with prompts.
4) Use "structural" deletions without prompts.
5) Use "any word" deletions without prompts.
6) Use "lexical" deletions without prompts.

Before discussing the rationale for sequence strategy IVB, several terms must be defined. The term "lexical deletions" refers to the deletion of such words as nouns, verbs, adjectives, and adverbs, which carry most of the substantive meaning in messages. The term "structural deletions" refers to the deletion of "structure" or "function" words such as determiners, subordinators, and verbal auxiliaries which, together with other clues, function as carriers of the structural or organizational meaning among words in a sentence. The term "any word deletions" designates the removal of both lexical and structural words. Sequence steps 1 and 2 are predicated upon the assumption that the two deletion types are order on a continuum from "concrete" to "abstract." The use of prompts (i.e., clues to the correct answer) is suggested because, despite this desirable continuum for learning, "lexical" items are so numerous in the language that they are difficult to predict. The introduction of "structural deletions" next is recommended because they represent a relatively more abstract category. Sequence step 3 provides practice with the help of prompts in taking the most common type of cloze test which is constructed by deleting both lexical and structural words. Steps 4, 5, and 6 proceed from the easiest to the most difficult type of deletions as indicated by previous research (Taylor, 1953). The structure words, although abstract, constitute a category with only a few members. This, with their high frequency of occurrence, makes them easy to predict. By the time the teacher reaches the fourth step, all prompts should have been "faded" so that the students can use only context clues in making predictions.

V. DETERMINING RESPONSE TYPES

The following section is concerned with the use of "fill-in-the-blank" and "multiple-choice" formats in construction of cloze exercises:

SEQUENCE STRATEGY VA

1) Use multiple-choice alternatives.
2) Use the "fill-in-the-blank" format.

SEQUENCE STRATEGY VB

In constructing multiple-choice options:

1) Use alternatives of *different* semantical meaning and from *different* grammatical classes than the correct choice.
2) Use alternatives of *different* semantical meaning and from the *same* grammatical class as the correct choice.

In suggesting sequence strategy VA, it was assumed that it is easier to use aided recall in making cloze responses than to fill in blank spaces. The use of only two alternatives is recommended due to space limitations within cloze exercises. Sequence strategy VB was based upon the assumption that the choice between two words such as "catch" and "book," (Step 1) for a particular blank space in a meaningful verbal sequence are easier than a choice between two words such as "catch" and "throw," (Step 2). Two other possible combinations for designing alternatives might be (a) words with the *same* semantical meaning but from a *different* grammatical class than the correct choice, and (b) words with the *same* semantical meaning and from the *same* grammatical class as the correct choice. The latter possibility would produce synonymous alternatives between which no choice could be made. The former possibility would seem to constitute a category without members (i.e., to change the grammatical class would necessarily change semantic meaning in this context).

VI. USING VISUAL CLUES

The teacher may wish to use visual prompts or clues to increase the probability of making successful responses on cloze exercises. If so, the following sequence listed in order from easy to difficult is suggested:

SEQUENCE STRATEGY VIA

1) Pictures
2) Letters
3) Underline marks
4) Length of blank space

Picture clues may be used with young children and adult illiterates. Either individual letters or letter groupings may be used as prompts at the beginning and/or end of a blank space. The number of underline marks in a blank space may indicate the number of letters required for that response. The relative length of a blank space is a more general clue for the desired word. (Prompts 2, 3, and 4 are only useful if exact word scoring is desired.) Of course, using combinations of any of these clues would reduce the probability of error more than would using any one clue alone. It should be noted that any clue or clues should be "faded" progressively so the reader becomes increasingly independent throughout the training period.

VII. USING REINFORCEMENTS

Cloze exercises lend themselves easily to various types of reinforcement procedures. This topic could be expanded to include a large amount of information about reinforcement techniques from the literature on behavior modification. However, for present purposes, three suggested instructional sequences will be given:

SEQUENCE STRATEGY VIIA

1) Use weak reinforcements for easy materials.
2) Use strong reinforcements for more difficult materials.

SEQUENCE STRATEGY VIIB

1) Use strong reinforcements for "uninteresting" materials.
2) Use weaker reinforcements as the task becomes self-reinforcing.

SEQUENCE STRATEGY VIIC

1) Use continuous reinforcement.
2) Use intermittent reinforcement.

It is assumed that the strength and weakness of reinforcements must be determined empirically for the individual(s) concerned and that the strength of reinforcement be proportional to the aversiveness of the task. The work on reinforcement schedules by B. F. Skinner and others will be helpful to the reader in applying sequence strategies VIIA, B, and C. Properly used, reinforcements can be employed to shape successive approximations to high level cloze performance and to maintain the strength of the desired cloze behaviors over prolonged time periods.

CONCLUSION

It should be emphasized that the mere mechanical administration of cloze materials as such cannot be expected to yield superior results in reading achievement. Indeed, it is somewhat surprising that many studies have shown such routine administration of cloze exercises yields at least as good results as traditional training procedures. In any case, these suggested sequence plans for teaching with cloze materials are only guidelines. The skillful teacher will try these sequences in various combinations and make careful observations to determine the best "mix." Even so, compared to the immense subjectivity and guesswork involved in determining most teaching sequences for improving reading comprehension, these specific operational strategies for sequencing instruction are relatively objective and carefully defined. As a technique for test construction, the cloze procedure has the advantage of providing precise step-by-step procedures for producing identical tests capable of being replicated independently by different test constructors. Movement toward such laboratory-type precision would be, in certain respects, highly desirable in the reading classroom for teaching purposes. It is hoped that the author's hypothesized sequence strategies will be a step in the direction of helping the cloze procedure to realize its potential for generating more precise and operationally defined procedures for teaching reading comprehension.

REFERENCES

Jongsma, E. *The cloze procedure as a teaching technique.* Newark, DE: International Reading Association, 1971.

Rankin, E. F., Jr. *An evaluation of the cloze procedure as a technique for measuring reading comprehension.* Unpublished doctoral dissertation, University of Michigan, 1957.

Rankin, E. F., Jr. The definition of reading comprehension. In Raygor, A. L. (Ed.), *College and adult reading* (Vol. 1). Minneapolis, MN: North Central Reading Association, 1962, 15-31.

Smith, F. *Understanding reading.* New York: Holt, Rinehart, and Winston, 1971.

Taylor, W. L. *The cloze procedure: How it predicts comprehension and intelligence of military personnel* (Technical Memorandum No. 13 to the United States Air Force). Urbana, IL: University of Illinois, Human Resources Research Institute, Division of Communications, 1953, 1-22.

ROGER KLUMB
NCS/Educational Systems

WAYNE OTTO
University of Wisconsin-Madison

DEVELOPING AN ESSENTIAL READING SKILLS PROGRAM FOR POST-ELEMENTARY STUDENTS

The quality of education — reading instruction in particular — has been subjected to continuous examination, evaluation and criticism in the past decade. While many of the resulting charges and counter charges have yet to be substantiated, two facts are well known. First, many students are graduated from high school without having attained functional reading ability. Second, many attempts to tackle this problem are resisted by both teachers and students, and, therefore, produce no measurable results. The authors' position is that these failures are largely attributable to a lack of focus in the goals that are set for secondary reading programs and to a general insensitivity of such programs to the realities of the organization of secondary schools and to the priorities of secondary school teachers. Our attempts to deal with these problems are described in this paper.

A COLLABORATIVE EFFORT

In the fall of 1974, the specific need for a program to improve the basic reading skills of post-elementary school students was identified by a reading improvement committee of administrators and teachers in the Duluth, Minnesota, Public Schools. The committee estimated that between fifteen and thirty percent of Duluth's high school graduates had serious deficiencies in reading. This estimate is similar to estimates in other school districts in the nation. Nevertheless, the committee noted that the reading skill mastery of Duluth's elementary pupils had improved in recent years and that the *Wisconsin Design for Reading Skill Development* (WDRSD) (Otto & Askov, 1974), a systematic, objective-based reading assessment and management system might provide a basis for improving reading instruction and skill mastery at the post-elementary school level.

The Duluth reading committee then teamed with the University of Wisconsin Research and Development Center (R & D Center) staff and personnel from NCS/Educational Systems, respectively the developers and the publishers of the WDRSD, to develop a reading program designed for post-elementary students who lack adequate development of the reading skills required to cope with "the basic demands of everyday living." Using the WDRSD as a model, the R & D Center staff drew up specifications for the post-elementary program. The program, *Design for Reading: Secondary* (DR:S), is based on the same assumptions as the WDRSD: (1) instructional objectives that describe essential basic reading skills can be identified and (2) those instructional objectives can provide a basis for reading assessment instruments and instructional materials. Component parts of the DR:S program include (1) a screening test, (2) teacher-directed reading activities, (3) skill-focused reading activities, and (4) criterion-referenced tests.

Specifications for the DR:S were submitted to NCS/Educational Systems for development. NCS/Educational Systems has worked closely with the R & D Center in field testing, publishing, and providing teacher inservice instruction during the development of the DR:S. Staff personnel from Duluth, the R & D Center and NCS have combined their efforts in planning, writing, pilot testing, and revising the program.

DEVELOPMENTAL RESPONSIBILITIES

The approach taken in developing the program derives focus from explicit objectives and pragmatic appeal from the collaborative efforts of a university research and development team, secondary school practitioners, and a publisher's staff.

The R & D Center staff took responsibility for identifying a list of basic skills and objectives and for assisting in the development and evaluation of (1) test items, directions, and manuals; (2) instructional strategies; and (3) a record-keeping system. In addition, the R & D Center staff assisted in the development of instructional materials.

The Duluth Public Schools staff assumed the following responsibilities: (1) review all materials prior to implementation; (2) develop a record-keeping system; (3) develop implementation strategies; and (4) provide subjects for pilot-testing assessment and instructional materials. The school people, then, were concerned mainly with matters of credibility and useability.

NCS/Educational Systems agreed to underwrite the developmental costs, to furnish copies of materials for the pilot test in Duluth, and to manage the development and production of the program. The NCS staff was accountable for assisting in the development and evaluation of (1) test items, directions, and manuals; (2) implementation strategies; (3) instructional materials; and (4) the record-keeping system.

Thus, the program is rooted in the realities of existing scholarship, the classroom, and availability in published form.

COMPONENTS OF THE PROGRAM

Skills and Objectives

The Design for Reading: Secondary is based on reading skills and objectives and is divided into two levels. Level I includes prereading skills (visual and aural), sight vocabulary (frequently used words), and word attack skills (phonics and structural analysis). Level II includes reading comprehension skills (words, sentences, selections and sequence) and reading resource skills (map, graph, table, and reference). The prereading skills and objectives are adapted from the *Pre-Reading Skills Program* (Venezky, Green & Leslie, 1975), another product developed at the Wisconsin R & D Center. The word attack, comprehension, and reading resource skills are adapted from the *Wisconsin Design for Reading Skill Development*. The sight vocabulary words used in the secondary program are the first 300 words of *The Great Atlantic and Pacific Word List* (Otto & Chester, 1972).

Assessment Materials

Placement Test. The purpose of the placement test is to assess students' reading skill mastery in the areas of word attack, comprehension and resource (map, graph, and table) skills. Level I of the placement test, the word attack portion, includes nonsense words and frequently used phonically and/or structurally regular real words. The Level I tests are designed to help teachers determine whether or not students have mastered selected consonant, vowel, and structural analysis skills. Level II of the placement test has two sections: comprehension and resource skills. The comprehension portion includes passages that help to determine whether or not students have mastered the skills needed (1) to use context clues; (2) to locate and comprehend details in sentences; (3) to identify the main idea of a selection; (4) to find relationships and conclusions in a selection; and (5) to comprehend sequence concepts. Short passages are used to assess context clues. Longer selections that have the necessary structure and information are used to assess details, main idea, and relationships and conclusions. The resource skills portion of the test focuses on a variety of skills required for interpretation of maps, graphs and tables.

The results of the placement test help teachers identify specific skills which need to be taught to individual students.

Criterion-Referenced Skill Tests. Criterion-referenced tests were constructed to assess individuals' mastery of the objective stated for each reading skill in the program. The tests are of two types: (1) group-administered paper-and-pencil tests and (2) individually-administered performance tests, both based on characteristics of the skills and the stated objectives.

The purpose of the criterion-referenced tests is to assess students' mastery of each skill/objective in the program. The tests can be used in implementing one of two instructional sequences. First, the tests can provide the basis for a teach-test sequence; that is, after students receive instruction in a particular skill, they can be tested for skill mastery. Second, the tests can provide the basis for fuller individualized skill instruction; that is, after students take the placement test and general skill deficiencies are identified, the criterion-referenced tests can be administered for diagnostic (pretest) purposes. Then an individualized instructional program based on each student's need can be implemented. Then, after appropriate instruction, each student can be retested to measure progress in reading skill development. This is the classic test-teach-test sequence.

Instructional Materials

The instructional materials developed for the *Design for Reading: Secondary* include teacher-directed activities and skill focused reading activities that are based on the program's objectives.

Teacher-Directed Activities. The teacher-directed activities are included in three separate teacher's handbooks: one handbook for prereading and word attack skills, one for reading comprehension skills, and one for the reading resource skills. Each teacher's handbook contains teacher-directed activities designed to develop individual student's reading skills. The teacher's handbooks were written and compiled to correspond to the sequence of skills in the *Design for Reading: Secondary*. For each skill, a handbook contains a teacher-directed learning activity section that lists (1) the skill name; (2) the instructional objective; (3) an elaboration of the objective; (4) suggestions for teacher-directed activities, e.g., transparency masters, chalkboard demonstrations, topics and examples for discussion, etc.; (5) examples for skill application, e.g., word lists; and (6) references to the appropriate skill tests.

The purpose of each teacher's handbook is to provide post-elementary teachers with (1) an understanding of the basic reading skills that are included, and (2) guidelines for teaching the basic reading skills. The lessons are designed to provide students with opportunities to practice skill application in isolation and in context. Words from the *A & P Word List* in combination with words used less frequently make up the sentences and selections which students read to apply the various comprehension skills. Both formal and informal procedures, based on the characteristics of the skill and objectives, are provided in the teacher-directed activities.

Skill-Focused Reading Activities. Various skill-focused reading activities are used to review and reinforce the reading skills presented in the teacher-directed activities. In the prereading and word attack area, the two major activities involve (1) memorization of certain sound/symbol-symbol/sound correspondences and phonic generalizations, and (2) application, in the context of printed words, of previously attained knowledge in new or different ways. The strategies used depend upon the inherent characteristics of each skill (objective) and the order in which the skills are taught. Word attack activities are organized and presented in three separate activity books: consonants, vowels, and structural analysis.

In the reading comprehension area, skill related instructional activities are also provided. The activities are consistent with the concept that reading is a process rather than a separate content area, i.e., that students may continue to improve their mastery of basic skills in increasingly more demanding contexts. Comprehension activities are organized and presented in four separate reading books: word meaning, sentence meaning, selection meaning, and sequence.

In the reading resources skill area two types of skill-focused activities are provided. In the first type, examples of maps, graphs, tables, and schedules that call for the application of specific skills are presented to the students. The activities require the students to solve problems by applying the skills. In the second type reference sources like dictionaries, newspapers and encyclopedias are presented along with suggestions for realistic projects. The activities provide opportunities to work with real resources on real problems.

FIELD TESTS OF PROGRAM COMPONENTS

All of the *Design II* materials have been field-tested by schools in Windsor,

113

Connecticut; Niagara Falls, New York; Milwaukee, Wisconsin; Jefferson County, Colorado; and Duluth, Minnesota. The purpose of the tryouts was to gather three types of information: (1) teachers' reviews of the materials, (2) students' test results to verify test item reliability, and (3) students' test results to get preliminary evidence of the impact of the skill-focused instruction.

The teachers' reviews ranged from subjective comments with regard to the maturity level and style of the material to specific recommendations related to implementing the materials in the classroom. The consensus was that these objective-based materials written specifically for post-elementary students were: (1) conducive to teaching basic reading skills to post-elementary students, (2) helpful in focusing the goals of the reading program, and (3) responsive to the organization of a post-elementary school and to the instructional priorities of post-elementary teachers.

Since the spring of 1975, the *Design II* tests have been critically reviewed twice by the Duluth reading committee. They were also critically reviewed and item analyzed by the University of Wisconsin Research and Development Center evaluation staff. Based on these reviews and the field-test results, the tests were revised twice.

All of the *Design II* tests are criterion-referenced. Basic reading skills were identified, and behavioral objectives were stated for each skill. All of the skills were then grouped into units, and the test items were written to assess the behavior stated for the objectives. That is, a correct response to each test item

Table 1

Proficiency Test Reliability Estimates

Test	Reliability Estimate
Word Attack	
Test 1a - Consonants	.522
Test 1b - Vowels	.803
Test 1c - Structural Analysis	.633
Word Attack Unit	.838
Reading Comprehension	
Test 2a - Word Meaning	.794
Test 2b - Sentence Meaning	.712
Test 2c - Selection Meaning	.655
Test 2d - Sequence	.564
Reading Comprehension Unit	.886
Reading Resources	
Test 3a - Maps	.787
Test 3b - Graphs and Tables	.741
Reading Resources Unit	.847

requires the student to demonstrate the behavior as stated in the objective. Content validity is apparent upon comparison of the test items with the corresponding objectives.

Statistical analysis was used to compute the reliability estimates of internal consistency. For example, in the Proficiency Test the reliability estimates ranged from .522 for the Consonants test to .886 for the Reading Comprehension Unit. Table 1 lists the reliability estimates for the Proficiency Test.

The reliability estimates for the individual skill tests were all higher than those on the Proficiency Test, which is probably due to the fact that the individual tests have more test items per skill than the subtests of the Proficiency Test.

At the present time, data are not available to calculate the impact of *Design II* on students' test results because the final edition of the teacher-directed activities and the skill-focused reading activities has not been completed. Completion is scheduled for January, 1978.

REFERENCES

Otto, W., & Askov, E. *The Wisconsin design for reading skill development: Rationale and guidelines.* Minneapolis: National Computer Systems, 1974.

Otto, W., & Chester, R. Sight words for beginning readers. *The Journal of Educational Research,* 1972, *65,* 435-443.

Venezky, R., Green, M., & Leslie, R. *Evaluation of the pre-reading skills program 1971-1974.* Madison: The University of Wisconsin Research and Development Center for Cognitive Learning, 1975.

JOYCE D. JOHNSTON
University of Georgia

IMPROVING HIGH SCHOOL/COLLEGE ARTICULATION FOR COMPENSATORY STUDENTS THROUGH EARLY BASIC SKILLS EVALUATION*

One of the problems for compensatory instructional programs at colleges with selective admissions policies is how to choose, from students rejected for regular admission, those with sufficient academic potential to make them likely candidates for success in the regular curriculum in an economically feasible time. A number of investigators have cited the need for both careful and early identification of students who could be expected to benefit from a compensatory program (Losak, 1972; Chadwick & Walters, 1975). Other workers have studied the specific relationship between the basic skill of reading and academic success in college (Humber, 1944; Pedrini & Pedrini, 1975).

Identification of entering college students who, while deficient in basic skills, do possess academic potential means fully utilizing high school information. Several authorities have stressed this mutual dependence of high school and college programs and, at the same time, expressed concern over problems of articulation between these two levels (Menacker, 1975; Wattenbarger, 1972). Moving from one level of the "educational ladder" to another is not always an efficient process. Coordination of testing programs is cited as a means of improving "efficiency and effectiveness" in education, qualities especially needful in compensatory programs if the students enrolled in such programs are going to catch up before they give (or are given) up.

This study was undertaken to determine relationships among high school achievement tests in the basic skills of reading, English, and mathematics and other predictors of likely college performance at the University of Georgia for students who were rejected for regular admission and subsequently enrolled in the compensatory program. One consideration was whether the tests given at the beginning of Grade 11 in all Georgia high schools through the Statewide Testing Program might be useful early identifiers of students who would be ineligible for regular admission at the University of Georgia two years later. In addition, various reading scores were further analyzed for their ability to predict how well students enrolled in the compensatory program performed in social science courses in the regular curriculum.

---

*This report is based on a doctoral dissertation completed at the University of Georgia, and the author wishes to acknowledge the assistance of her co-major professors, Drs. George E. Mason and Robert A. Palmatier.

METHOD

The subjects were 98 students in the Special Studies Division at the University of Georgia during summer quarter, 1975. The selection of the sample from the total group of 206 was based primarily upon the availability of the high school achievement test data. Graduates of 37 different high schools located throughout Georgia were included.

The instruments used were the reading, English, and mathematics subtests of the *Test of Academic Progress* (TAP) given at the beginning of Grade 11 in Georgia high schools; the SAT, verbal and math, given in Grade 12; and the reading, English, and mathematics subtests of the *Comparative Guidance and Placement Test* (CGP), given at college entrance. The CGP is the placement test required by the University System of Georgia for students in the Special Studies Program in any institution in the System. Other variables included were the high school grade average (HSA) and the grade point average (GPA) in social science courses over a two-quarter period. The data were collected from University of Georgia records.

Pearson product-moment correlation coefficients were computed for all combinations, using the .05 level as the criterion of significance. McNemar's t-test for the significance of the difference between non-independent correlations was used to determine whether the difference between any two correlation coefficients was significant. To study the predictive ability of the reading-related measures, multiple regression analysis was used, utilizing a computer program that entered all possible sequential orders of the variables.

RESULTS AND DISCUSSION

A summary of the correlation is found in Table 1. For the students in this sample, achievement in reading, English, and math at the beginning of Grade 11 (TAP scores) appeared to have a significant, positive relationship to college aptitude (SAT scores) and to freshman placement in college (CGP scores). The magnitude of the relationships for reading and math (r = .50 or above) was greater than for English, however (r = .21 and .31, respectively). Whether either the SAT scores or the CGP scores were more highly related to the high school achievement (TAP) scores could not be determined since the t values were nonsignificant.

The correlations between high school achievement in Grade 11 (TAP) and high school grade average (HSA) two years later, however, presented a different picture. Eight of the ten correlations proved to be negative. The direction of the correlations between the standardized test scores and the more subjective measure of grade averages is important since it is the opposite of what is usually seen. The difference between the correlations between the total high school achievement score (TAP-Total) and the SAT-Total about a year later (r = .58) and between the TAP-Total and the high school grade average (t = -.16) was significant (t = 5.24, $p < .001$). Only about six months elapsed between the time the SAT was taken and the final grade average period in the spring of 1975. Such a short intervening time period seems inadequate for enough educational growth to explain that much difference among variables assumed to be highly related. Furthermore, the correlation between SAT-Total and HSA was also negative and moderately strong ($r = -.47, p = .0001$).

Table 1

Pearson Product-Movement Coefficients of Correlation Between Variables in Study[a]

	TAP-E	TAP-M	CGP-R	CGP-E	CGP-M	SAT-V	SAT-M	TAP Total	SAT Total	HSA
TAP-R	.46 (.0001)	.26 (.0094)	.50 (.0001)	.34 (.0009)	.35 (.0006)	.54 (.0001)	.20 (.0489)	.77 (.0001)	.53 (.0001)	-.20 (.0475)
TAP-E	—	.27 (.0072)	.24 (.0159)	.31 (.0020)	.16 (.1079)	.21 (.0364)	.19 (.0539)	.77 (.0001)	.29 (.0039)	.10 (.6622)
TAP-M	—	—	.14 (.1736)	.11 (.2814)	.55 (.0001)	.08 (.5815)	.55 (.0001)	.69 (.0001)	.46 (.0001)	-.25 (.0126)
CGP-R	—	—	—	.45 (.0001)	.20 (.0512)	.63 (.0001)	.10 (.6509)	.39 (.0002)	.53 (.0001)	-.18 (.0657)
CGP-E	—	—	—	—	.21 (.0382)	.38 (.0003)	.11 (.2667)	.34 (.0009)	.36 (.0006)	.04 (.7066)
CGP-M	—	—	—	—	—	.08 (.5341)	.63 (.0001)	.48 (.0001)	.51 (.0001)	-.26 (.0096)
SAT-V	—	—	—	—	—	—	-.05 (.63[illegible])	.37 (.0004)	.70 (.0001)	-.33 (.0012)
SAT-M	—	—	—	—	—	—	—	.42 (.0001)	.68 (.0001)	-.32 (.0016)
TAP-Total	—	—	—	—	—	—	—	—	.58 (.0001)	-.16 (.1152)
SAT-Total	—	—	—	—	—	—	—	—	—	-.47 (.0001)

[a]Probability listed in parenthesis below correlations.

N = 98

Key:
TAP-R = Test of Academic Progress, Reading Subtest
TAP-E = Test of Academic Progress, English Subtest
TAP-M = Test of Academic Progress, Math Subtest
CGP-R = Comparative Guidance & Placement Test, Reading Subtest
CGP-E = Comparative Guidance & Placement Test, English Subtest
CGP-M = Comparative Guidance & Placement Test, Math Subtest
SAT-V = Scholastic Aptitude Test, Verbal
SAT-M = Scholastic Aptitude Test, Math
TAP-Total = Test of Academic Progress, Total Score
SAT-Total = Scholastic Aptitude Test, Total Score
HSA = High school average (grades)

The negative direction appears to reflect the atypical nature of these students, since an earlier study (Irvine, 1966), which used regularly admitted University of Georgia students, found the relationships between HSA and SAT to be of similar magnitude but of opposite direction. Irvine's HSA/SAT-V correlation was .23, compared to -.33 in the current study. His HSA/SAT-M correlation was .30, compared to -.32 in the study.

Why there were apparent discrepancies between grade averages and SAT scores for the compensatory student sample is not clear. It may be straightforward grade inflation. (Grades are, after all, the more subjective measure.) It could also be that the student who is perceived by a teacher as having less academic aptitude but as very industrious is given a higher grade than is actually earned, while the student thought to be able but lazy gets only the legal minimum, so to speak.

The above results suggest that the earlier high school test (TAP) may well be a useful added indicator of the deficiencies that ultimately place students in compensatory programs, after being rejected for regular college admission, *despite* their high school grade averages. This conclusion was strengthened by the analysis of the predictive ability of the reading scores in the study. The three predictor variables for reading were the Grade 11 reading achievement subtests (TAP-R), the SAT-V, and the Special Studies reading placement test (CGP-R). The criterion variable was GPA in regular college social science courses. This analysis is summarized in Table 2.

Table 2

Prediction of Social Science GPA from Reading-related Tests

Predictors	Correlation with Criterion (GPA)[a]	Squared Correlation (Variance)	F value for Full vs. Partial Model
SAT-V	.16	.03	6.66*
CGP-R	.30*	.09	3.30*
TAP-R	.35*	.12	1.66
CGP-R + SAT-V	.31*	.09	6.51*
TAP-R + SAT-V	.35*	.12	3.32
TAP-R + CGP-R	.38*	.15	0.98
TAP-R + CGP-R + SAT-V	.40*	.16	

[a] For single predictor, correlation is zero-order r; with more than one predictor, correlation is multiple R.

*$p < .05$
$N = 90$

No significant differences were found between the full model, using all three predictors, and either of the two predictor models that included the TAP-R score, or between the full model and the TAP-R used as a single predictor. That is, using TAP-R alone to predict social science GPA for this sample did not appear to be significantly different from using a combination of all three test scores. Only when TAP-R was dropped from the predictor set was there a significant difference between the full and partial models. Thus, the high school reading test

(TAP-R) was a slightly better predictor of performance in a regular college course than either the SAT-V for CGP-R (or, in fact, all three reading tests combined), and yet the TAP-score is not included in the decision-making input for compensatory program placement.

While this study was not sufficiently broad to make firm generalizations, it does support the idea that it would be worthwhile for colleges in Georgia to use data from the Statewide Testing Program in high school for help in decision-making in their college compensatory programs. For the compensatory program students in this sample, the high school grade average as an indicator of college success was not as reliable as was expected. Follow-up studies are needed to determine what indicators are the most reliable for atypical students. The results here suggest that reading ability may be a possibility.

If the relationships found here could be replicated at other colleges in the University System, the role of the Georgia high school achievement test as an early identifier of students weak in basic skills could be established. This opens up the possibility of closer high school/college cooperation, with compensatory intervention beginning earlier — in the high schools. The goal of such coordination would be to reduce the number of students needing remedial help when they enter college.

REFERENCES

Chadwick, R., & Walters, H. *Reading relationship study: Success/failure rate of first semester college students identified as poor readers.* Sacramento, Calif.: Consumnes River College, 1975. (ERIC Document Reproduction Service No. ED 103 072)

Humber, W. J. The relationship between reading efficiency and academic success in selected university curricula. *Journal of Educational Psychology,* 1944, *35,* 17-26.

Losak, J. Do remedial programs really work? *Personnel and Guidance Journal,* 1972, *50,* 383-386.

Menacker, J. *From school to college: Articulation and transfer.* Washington, D.C.: American Council on Education, 1975.

Pedrini, D. T., & Pedrini, B. C. Reading ability and college grades. *College Student Journal,* 1975, *9,* 37-42.

Wattenbarger, J. L. Articulation with high schools and four-year colleges and universities. In T. O'Banion, & A. Thurston (Eds.), *Student development programs in the community junior college.* Englewood Cliffs, N.J.: Prentice-Hall, 1972.

ELEANOR HABURTON
Valencia Community College, Orlando, Florida

IMPACT OF AN EXPERIMENTAL READING-STUDY SKILLS COURSE ON HIGH-RISK STUDENT SUCCESS IN A COMMUNITY COLLEGE

In examining the implications of the mission statement, "to maximize the potential of all students who wish to enter", recent research studies in open-door community colleges have focused on several aspects of reading instruction for high-risk students. Variables studied include grade point average (GPA), attrition, persistence, self-concept, and student participation in reading programs with content-area emphasis.

Roueche and Kirk (1973), in a study of programs for high-risk students in five selected community colleges, found such programs helped students to attain higher GPAs and to persist in college.

Astin (1975), in his 1968-1972 study of student persistence, stated that GPA was more closely associated with persistence than any other variable, and concluded that anything that can be done to enhance academic performance will also tend to reduce attrition rates. Cohen and Brawer (1970) reported that withdrawal was related to financial pressure and to low grades.

Clarke (1975) reported that the community college has had some success in reaching the high-risk student as seen by increasing enrollment of these students and by the increasing number of graduates; however, large numbers are not succeeding, and the attrition rate is decreasing only slightly.

Investigations of student self-concept and of methods of determining individual cognitive styles to augment learning were carried out by Lang (1972), Nunney and Hill (1972), Kahn (1974), and Tillman (1976). Research conclusions were that learning styles and attitudes toward self are highly significant factors in academic success.

Several studies advocated teaching of reading and study skills directly with subject or content-area material. Beitler and Martin (1972) successfully worked with vocationally oriented students in a learning center aimed toward mastery of the content materials of career programs. Soll (1972) taught study skills by using subject assignments rather than in isolation in traditional programs using commercial materials or mechanized labs.

The purpose of the present study was to investigate the impact of a reading-study skills course on the academic success of a group of high-risk community college students in terms of: (a) grade point average, (b) number of course withdrawals, (c) attrition (dropping out of school) during the semester and between-semesters, (d) number of course hours successfully completed, (e) self-reported posttest scores on study habits and attitudes as measured by the *Survey* *of Study Habits and Attitudes* (SSHA), and (f) self-reported posttest scores on self-concept as measured by the *Tennessee Self Concept Scale* (TSCS). Also, the two experimental classes were compared on the hypothesized variables to determine if the teacher variable made a difference.

METHOD

Subjects

The 36 subjects comprising the experimental group were those who voluntarily enrolled in two reading-study skills classes of 15 and 21 during the fall semester, 1975, at Valencia Community College at Orlando, Florida. The experimental classes were taught by two reading instructors.

The 36 control group students also volunteered to enroll in the reading classes but were prevented from doing so either because the classes were closed or because of schedule conflicts. The control group was enrolled in regular classes and received no treatment in reading during the fall semester.

The groups were matched on the following variables: (a) voluntary participation in a previous basic reading course, (b) age, (c) sex, (d) race, (e) pretest reading comprehension score below the 30th percentile on the *Iowa Silent Reading Test, Form E*, and (f) carrying a minimum class load of 9 hours.

Of the 72 total participants, there were 46 black students, 26 male and 20 female. There were 20 white male and 6 white female students. The breakdown in age groups was as follows: (17-19) — 30; (20-25) — 18; (26-39) — 18; (above 40) — 6.

Procedure and Materials

The experimental program was one semester in length with classes meeting either three days a week for 50 minute periods or two days a week with sessions of 1 hour and 15 minutes each for 14 weeks. In addition, two individual conference periods per week were arranged with the instructors for each student as needed.

The plan of the course was to maximize chances for academic success in content courses by (a) offering three hours of transfer credit for the reading class, (b) focusing on "survival" skills needed in content area courses and providing for direct transferal, (c) providing for self-pacing, (d) utilizing group instruction with group interaction as a motivational factor, (e) accommodating various learning styles through multi-media presentation, (f) evaluating progress by mastery tests, and (g) attempting to enhance self-concept by student orientation of learning, teacher attitude and individual conferences.

Materials used included the students' textbooks and an adaptation of the textbook *Probe* and the 30 accompanying cassette tapes developed by Glock, Bender and Dennis (1975).

The posttests for study habits and self concept were given to the experimental group in their classrooms. The control group members were tested individually by appointment in the reading lab since they were not enrolled in reading classes.

RESULTS

To test the 8 hypotheses, t tests and chi-square tests were performed on the data. On the basis of statistical analysis, it was concluded that the experimental group, ($M = 2.50$, $SD = .63$), received significantly higher GPAs than the control group, ($M = 1.89$, $SD = .64$), $t(58) = 3.71$, $p < .01$. To overcome possible grading bias, the GPA was also determined by eliminating the reading course grades from the total GPA scores of the experimental group; however, the GPA without the reading grade was still significantly higher for the experimental group, ($M = 2.48$, $SD = .68$), $t(58) = 3.37$, $p < .01$.

The experimental group, ($M = 3.56$, $SD = 3.63$), withdrew from fewer course hours than the control group, ($M = 5.72$, $SD = 3.89$), $t(70) = 2.44$, $p < .05$. Also, the experimental group, ($M = 7.89$, $SD = 4.37$), successfully completed more course hours than the control, ($M = 5.58$, $SD = 4.28$), $t(70) = 2.26$, $p < .05$. Since the number of course hours in which the students enrolled varied from a minimum of 9 up to 19, a t test was run on this variable to determine if there was a significant difference in course loads undertaken at the beginning of the semester. There was no difference: Experimental ($M = 12.42$, $SD = 2.14$), Control ($M = 12.50$, $SD = 1.96$), $t(70) = .17$, $p > .05$.

A chi square comparison showed no significant difference between the groups on attrition during the semester. A total of 33 experimental and 34 control remained.

On six of the seven subtests of the *Survey of Study Habits and Attitudes* the experimental group scored significantly higher than the control (see Table 1). On one subtest, Educational Acceptance, there was no significant difference.

Table 1

Experimental versus Control Comparisons for
the Survey of Habits and Attitudes

	Experimental		Control		*t* Values
Variable	*M*	*SD*	*M*	*SD*	
Delay Avoidance	22.83	9.41	17.00	6.54	2.83*
Work Methods	25.40	8.59	18.64	7.28	3.35*
Study Habits	48.50	16.76	35.70	12.14	3.44*
Teacher Acceptance	31.40	8.44	25.12	9.11	2.84*
Study Organization	107.27	26.97	85.28	24.15	3.40*
Study Attitudes	58.77	13.55	49.58	15.20	2.54**
Educational Acceptance	27.37	7.54	24.15	7.58	1.69

* = significant at .01
** = significant at .05

There was no significant difference between the groups on self-reported scores on self-concept as measured by the 14 subtests of the *Tennessee Self Concept Scale*. Representative of these findings is the score on the total positive (TP) test, the most important single score in the measure which reflects the overall level of self-esteem. The data for this subtest are: Experimental ($M = 332.73$, $SD = 31.76$) versus control ($M = 343.69$, $SD = 32.68$), $t(60) = 1.34$, $p > .05$.

Also, there was no statistically significant difference on any of the variables between the two experimental reading classes, and thus it was concluded that the teacher variable did not bias the results.

DISCUSSION

The study indicates that high-risk students who take long term reading improvement classes tend to succeed better in academic courses than do like students who do not take such programs.

The higher GPAs and the higher self-report scores on the SSHA may be due in part to the fact the experimental group learned workable study habits and became more sophisticated in surviving in the college situation. The self reports on the TSCS were in some cases quite unrealistic (by both groups), and this measure would perhaps be more useful for counseling individual students than for evaluating self-concept changes.

CONCLUSIONS

The focus of the study was to synthesize as many as possible of the research-based practices in reading program content and management that have been reported to date as holding promise for the success of high-risk students in academic courses. The findings generally support the results of related studies that have reported the success of reading instruction when it was directly related to the students' academic courses.

In contrast to earlier research reported in the 1960s which often concluded that remedial reading programs for high-risk community college students were unsuccessful (Losak, 1970), the present study lends some support to the more optimistic research of the mid-1970s (Roueche, 1976) in which exemplary practices seem to yield a higher success factor.

Since there is increasing evidence that many high-risk students can succeed over time, there is a need for longitudinal studies of such students who are involved in long term (two to three year) programs to improve their communications skills. Such programs need to be cooperatively planned with the content-area instructors so the pragmatic student can move forward toward his academic or career goal while working to overcome his language skill gaps.

REFERENCES

Astin, A. W. *Preventing students from dropping out.* San Francisco: Jossey-Bass, 1975.

Beitler, L., & Martin, I. A learning center for career-oriented students. *Twenty-first Yearbook of the National Reading Conference II.* Boone, N. C.: National Reading Conference, Inc., 1972.

124

Brown, W. F., & Holtzman, W. H. *Survey of study habits and attitudes manual, Form C.* New York: The Psychological Corporation, 1965.

Clarke, J. R. *Commitment to the nontraditional student* (Topical Paper No. 51). Los Angeles: ERIC Clearinghouse for Junior Colleges, University of California, 1975.

Cohen, A. M., & Brawer, F. B. *Student characteristics: Personality and dropout propensity* (Monograph series No. 9. ERIC Clearinghouse for Junior Colleges). Washington, D. C.: American Association of Junior Colleges, 1970.

Farr, R. *Iowa silent reading test, manual of directions.* New York: Harcourt Brace Jovanovich, Inc., 1973.

Fitts, W. H. *Tennessee self concept scale manual, counselor form.* Nashville: Counselor Recordings and Tests, 1965.

Glock, M.D., Bender, D., & Dennis, A. *Probe.* Columbus, Ohio: Charles E. Merrill, 1975.

Kahn, G. B. Expectations of open-enrollment students in the community college reading program. *Twenty-third Yearbook of the National Reading Conference.* Clemson, S. C.: National Reading Conference, Inc., 1974.

Lang, C. M. A study of the effects on learning of matching cognitive styles of students in nursing education. Doctoral dissertation, Michigan State University, 1972.

Losak, J. Research and compensatory education: What are we doing? *Florida Community-Junior College Inter-Institutional Research Council Publication* (University of Florida, College of Education), January 1970, 3-10.

Nunney, D., & Hill, J. *Personalized educational programs.* Bloomfield Hills, Mich.: Oakland Community College Press, 1972.

Roueche, J.E. Creating an environment for learning. *Community and Junior College Journal,* 1976, *46,* 48-50.

Roueche, J. E., & Kirk, R. W. *Catching up: Remedial education.* San Francisco: Jossey-Bass, 1973.

Soll, L. Learning by doing. *Journal of Reading,* 1972, *15,* 496-499.

Tillman, C. E. Personality types and reading gain for Upward Bound students. *Journal of Reading,* 1976, *19,* 302-306.

ANN MARIE BERNAZZA HAASE
University of Arizona

SHIRLEY DeSHIELDS
University of Massachusetts

MERVIN D. LYNCH
Northeastern University

EVALUATING THE COLLEGE READING PROGRAM: A MANAGEMENT AUDIT

Much has been written in the field of college and adult reading on curriculum and programs for readers in the college and adult population. Seldom, however, is the administration of such programs discussed. Usually, the act of administration is seen as a secondary phenomenon worthy of little cognizance and less respect. This lack of respect may indicate the act of administration, i.e., the act of service as a body politic, as a task which anyone can do without much effort. Yet, is underplays the value of the service by disregarding its management function. This function, whether in industry or education, consists of taking charge, directing, and guiding the particular circumstances in which it operates, so that those being administrated can perform their task with little interference.

To academics, the word management is distasteful. Academics believe themselves capable and possessive of the right to manage their own affairs, professional as well as personal. Administrative interference in faculty affairs is considered taboo. Yet, faculty readily yield the administrative function of a reading clinic or department to a colleague. In this act, they relinquish their right to management — to control and direct what is to happen. Academics state they make the decisions; the administrator carries them through. If so, then, what is the value of administrative management in reading clinics and reading departments? It is the purpose of this paper to provide a limited answer to that question.

MANAGEMENT ANALYSIS

A management analysis was performed on a reading skills program in a large northeastern university during the 1975-76 academic year. The program consisted of one administrator who reported to the head of the rhetoric program who in turn reported to an associate provost in charge of special programs. The reading program staff consisted of two administrative aides, one visiting faculty member, and 16 part-time doctoral students, all of whom taught in the program. The program serviced about 300 students per semester.

The analysis examined three areas: two internal and one external. The internal factors included a *who* and a *what* component. *Who* did what, when, where, and why? *What* were the functions and processes that occurred, such as grading and transfer of students in and out of the program? The external factor focussed upon outside influences and/or events that impinged upon the reading program.

Interview schedules were used to collect information on each of these three factors. Included were observations by the evaluator.

Role of Administrative Manager

Viewpoint. The director of the reading program was asked for a point of view on administration. She stated: "An administrator's job is to make sure the reading clinic staff does the best job possible in helping students learn to read effectively." The director added that much depended on the nature of the staff and she took special pains to insure its quality.

In response to a question about budgeting, the director stated the budget was crucial to the reading program. In fact, the director felt the budget could control reading curriculum development in that it either enhanced or restricted reading material acquisition, staff employment and development, and evaluation efforts.

The director also advised that budgeting made staff utilization look closely at amount of time spent at work by staff and that practices were modified in light of this examination. For example, during the 1974-75 academic year, the director analyzed the number of hours worked by half-time and quarter-time staff. She found half-time staff gave twelve hours of work for their twenty hours of paid work while quarter-time staff gave nine hours of work for their ten hours of paid work. From these data, the director decided to hire sixteen quarter-time staff instead of eight half-time staff for the 1975-76 academic year.

Responsibilities. The administrator in the reading program was responsible for the development and evaluation of the program, for staff employment and development, and was required to prepare and file reports requested by the director of rhetoric and to prepare annual budgets based on the zero based budgeting concept. The administrator also taught two three credit courses, one in reading and one in study skills.

Function and Processes

By convention and practice, the first two weeks of fall and spring semester in the reading center were primarily devoted to scheduling and assisting staff in accommodating to larger than expected classes and a greater number of independent study students than anticipated. Included also were tasks of material distribution and registration of late students. During the next three weeks of both semesters maintenance functions were established. This was necessary since there was about one-third staff turnover each semester. Staff turnover occurred because of graduation, academic withdrawal, acceptance of a better position elsewhere or a realization that teaching was not for that staff member.

During the remainder of both semesters, administrative and reading curriculum meetings were held by the administrator and staff. Administrative meetings were held by the administrator and staff. Administrative meetings consisted of information delivery, and presentation and discussion of current reading experiments, projects, and problems.

External Events

Crises. In addition to the responsibilities and functions inherent in the reading program, the administrator responded to two external events: denial of university credit for reading program courses and budget restrictions. The impact of these events on the staff caused severe morale problems which were alleviated by the director's immediate action of reassuring the staff the problem was administrative and not theirs and by negotiating with the head of rhetoric and the associate provost. These crises appeared to be the direct result of miscommunication between the head of rhetoric and the director of the reading program and the practiced duplicitousness of the head of rhetoric.

Staff Views of Administration. In order to understand how the staff view administration, they were asked specific questions concerning how they felt about administration generally and how they felt about their specific administrator. Sixteen of the eighteen saw administration in negative terms. Comments were made such as, "It's a joke!", and "a dirty job". When they were asked about how they felt about their director's performance on the job all eighteen saw her work as positive and fruitful in that the director provided an atmosphere for professional growth. Some comments were: "The director gives me space." "No hassles." "We can get our work done."

Thus, it appears that administration as an abstract concept is viewed negatively by this group; as a concrete personal concept, it is regarded positively.

CONCLUSION

The value of administration in this limited instance appears to lie in its service and management function in relation to its constituents. Administration, in this reading program, emphasizes organization and maintenance so that the staff can perform effectively, without disturbance. Also, it seems that competency of the administrator to perform her duties provides a sense of protection for the staff as well as determines the morale of the group.

Further research needs to be done in this area to determine the generalizability of these conclusions and to determine the impact of administration and administrative styles on the development and continuance of college and adult reading programs. Subsequent analysis might support the saying that as the administration goes, so goes the reading program.

HOMER COKER
JEFFREY L. LORENTZ
West Georgia College

GROWTH IN READING AS A CORRELATE OF PUPIL CLASSROOM BEHAVIOR

Systematic classroom observation instruments have, in recent years, been refined to the point where serious investigators can use them in studies of teaching effectiveness. A number of observation instruments are used in the Carroll County, Georgia, Competency Based Teacher Certification (CBTC) Project. In this Project, data were collected on more than 1800 pupils in 60 classrooms (grades 1-12) during the 1974-75 school year, and on more than 1300 pupils in 43 classrooms during the 1975-76 school year.

In an earlier study (Coker and Lorentz, 1975), the investigators reported significant relationships between classroom mean student Coping Style, teacher control, and Reading achievement. The present paper is a further examination of data collected as part of the CBTC Project. It was the purpose of the present study to examine the relationships between observed classroom behavior, specifically student Coping Style and pupil growth in Reading. In the present study, data on individual pupils, rather than on classroom averages were analyzed.

Several other recent studies have examined the relationship between observed classroom behaviors and pupil growth. Lahaderne (1968), Cobb (1972), and Samuels and Turnure (1974) have all found significant relationships between pupil attentiveness and achievement. McKinney and others (1975) used a modified version of Spaulding's CASES to investigate the relationships between classroom behavior patterns and academic achievement and found that combinations of discrete behaviors predicted achievement as well as IQ test scores and that the combination of test information and behavioral information predicted school year achievement better than was predicted by either alone.

In Follow Through studies, Soar (1973) found that a number of classroom behavior variables, including teacher control and structure, were related to increased pupil cognitive growth. Finally, Camp and Zimet (1974) found that positive classroom behavior reported by teachers, the lack of off-task behavior reported by observers, mental age, and measures of reading achievement were highly interrelated.

METHOD

Subjects

The subjects of the present study, 27 classroom teachers, grades 3-8, are a subset of the population of 60 teachers who participated in the CBTC Project during the 1974-75 school year. Only those teachers who taught Reading as a major subject are included in the sample selected for the present study.

Measures

The *Iowa Tests of Basic Skills,* Form 5 and 6, Levels Edition, (ITBS) (A. N. Hieronymus and E. F. Lindquist, 1971) was given to all pupils in grades 3 through 7 who were involved in the CBTC Project. The *Reading Comprehension* subtest only was used in the present study. The reported reliability and validity for the ITBS are adequate for use in the present study. In addition, the State of Georgia makes use of the ITBS Battery in the 4th and 8th grades in the state-wide testing program.

The Coping Analysis Schedule for Educational Settings (CASES), (Spaulding, 1970) is a category system designed to measure pupil socialization and consists of 19 categories of student "coping" behaviors which are identified by descriptive statements such as "Aggressive Behavior", "Self-directed Activity", "Observing Passively", etc.

An estimate of the socioeconomic status (SES) was obtained for each pupil based upon the occupation of the father (or the mother if she were the family's sole support). This occupational information was transformed to a 1-8 scale (1 = high status, 8 = low status) using a modified version of Warner, Meeker and Eells (1960) classification of occupations and levels. Pupils with missing SES were assigned the mean for their classroom, rounded to the nearest whole number.

Data Collection

The pretests were administered early in the Fall of 1974 in the classrooms of the 27 Carroll County teachers, and posttests administered again in the classrooms of the same 27 teachers in May, 1975. A group of specially-trained testers administered the tests in each classroom.

After the pretests had been administered, each of the approximately 600 pupils in the 27 classes was observed on two separate visits by observers using the CASES instrument. At the end of the year, two observations were again recorded on every pupil.

All CASES observations were made by experienced classroom teachers who had been employed full-time by the CBTC Project as observers and been especially trained by the author of CASES, Dr. Robert Spaulding. Observer agreement studies, conducted in non-project classrooms, confirmed that the observers were consistently able to maintain agreement of 80 percent or better.

Scoring

Tests were hand or machine scored, keypunched and verified. Observation data were keypunched from the data collection forms and verified.

ITBS scores were converted from raw scores to Standard Scores using the published Standard Score Tables (Houghton Mifflin, 1973). These scores used a normalized standard score scale with a mean of 80 and a standard deviation of 20 for the entire grade range, 3-8 data in a single analysis.

CASES frequencies for each pupil were scored using Spaulding's procedure (Spaulding, personal communication) to identify eight Coping Styles for each pupil as follows:

Style	Description
A	Aggressive, Manipulative
B	Inappropriately Self-directed
C	Passive, withdrawn
D	Peer Dependent
E	Compliant
F	Social, productive
G	Inner-directed, Task-oriented
H	Other-directed, Task-oriented
Overall	Composite of A - H

The CASES Coping Style scores for an individual pupil represent a dynamically-interrelated scale on which a predominant (i.e., highest) score in one or more styles is accompanied by low scores in the other styles. When a predominant Coping Style changes, another coping style will generally erupt as predominant.

When CASES observations were made, observers distinguished two types of classroom settings: *Teacher Directed* (TD) and *Program Directed* (PD).

A Teacher Directed setting is one in which the teacher is generally in charge of the group as a central, directing figure. A Program Directed setting is one in which the teacher is not central but rather, the pupils are generally working on their own in individual tasks (workbooks, etc.). There is evidence from Spaulding's studies and from an unpublished CBTC Project correlational study (1974) that changes in the classroom setting resulted in changes in the student's CASES scores. Scores were therefore obtained for students for the appropriate setting by combining data from both Fall visits only when collected in the same setting. The same procedure was followed for the CASES data collected in Spring visits. Thus for one pupil, one or two sets of pre-CASES scores could be obtained (e.g., both observations occurred in TD setting; both occurred in PD setting; one occurred in PD setting, and one occurred in TD setting), and one or two sets of post-CASES scores obtained.

Data Reduction

Pupils were sorted into Teacher Directed and Program Directed sub-samples only if Fall and Spring Reading scores as well as Fall and Spring Coping Style scores, along with SES, were available. If any of these variables were missing, that pupil was dropped from the sub-sample.

Changes in student Coping Style were computed in two ways. First, the difference between each of the Fall Coping Style scores and its corresponding Spring score was obtained as a measure of absolute change within each style. Second, since the change in predominant style may either be from one of the eight styles to another of the eight styles, or to the same style (i.e., the style remains the same), the change may be reflected as a category or cell in an 8 x 8 matrix.

Each of these 64 cells can be "scored" for each pupil by assigning an arbitrary value (i.e., *one*) to the cell representing the combination of Fall and Spring predominant Coping Styles and some other value (i.e., zero) to all the other cells. This procedure is known as "Dummy Variable" coding (Draper and Smith, pp.

134-135) and is a useful technique in regression analysis. Pupils with tied Fall or Spring predominant Coping Styles were dropped from this analysis.

The final sub-samples contained 342 pupils in the TD group and 233 in the PD group. Mean grade levels are very similar (TD= 5.5, PD= 5.1) while mean SES is the same (4.8) for both groups. Mean Reading achievement scores, however, differ (TD Fall= 74.2, PD Fall= 67.8; TD Spring= 75.8, PD Spring= 70.2). Since the groups were divided on the basis of CASES data and some pupils may appear with the same test scores in both groups, these differences in mean test scores are simply due to sampling error and not a result of classroom setting.

The TD and PD sub-samples have similar mean Coping Styles. The biggest difference between groups appears in Styles G (Inner-directed, task oriented) and H (Other directed, task oriented). The Teacher Directed group exhibits a much higher Style H score, while the Program Directed group exhibits a higher Style G score (TD Style H= 1.8-2.0, PD Style H=0.8-0.9; TD Style G= 0.1-0.2, PD Style G= 1.0-1.2). These two styles are the only *visible* means in the matrix (i.e., greater than one) for the groups.

Regression Analyses

To examine the relationships between the CASES Styles and Reading Achievement, a series of multiple regression analyses were carried out using the Biomedical stepwise multiple regression program BMD02R (Dixon, 1974). The dependent variable was always permitted to enter the equation in the order in which it accounted for additional proportions of the variance.

Four pairs of analyses of the TD and PD samples used Fall Reading, Grade, SES, and the Coping Styles as independent variables. Fall Coping Styles were entered in one set of analyses, Spring Coping Styles were entered in another; Change from Fall to Spring Coping Styles were entered in a third set of analyses, and in the fourth, the Fall, Spring, and Changes were all entered.

One additional pair of analyses used Fall Reading, Grade, SES, and the 64 dummy variables as independent variables.

RESULTS

In each of the analyses, the two best predictors of Spring Reading achievement were Fall Reading Achievement and grade level. Together these accounted for almost half of the variance in individual achievement (Multiple R = .64-.70).

However, once Fall Reading and grade level entered the equation, only two or three additional variables were significant ($p < .05$) increments to the equation. Table I presents the first six predictors from the analyses of Fall, Spring, and Changes in Coping Style.

The stepwise procedure will, if permitted, allow more variables to enter the equation than are statistically significant. If the number of independent variables is small in relation to the number of subjects, all the variables may enter the equation. The task of selecting the best equation from the many resulting from the stepwise procedures generally requires some personal judgment (Draper and Smith, p. 163). An examination of the summary tables revealed very little improvement in multiple R resulted from the addition of the sixth and subsequent

Table I

First Six Predictors of Spring Reading Achievement
Using Student Coping Styles

Variable	Multiple R	Multiple RSQ	F to Enter	Prob
Teacher Directed Setting				
Fall Reading	.62	.38	213.02	.00
Grade	.64	.41	14.90	.00
Fall Style H	.65	.42	3.78	.05
Fall Style E	.65	.43	5.05	.03
SES	.66	.43	2.41	.12
Decrease in Style B	.66	.43	1.74	.19
Program Directed Setting				
Fall Reading	.67	.45	185.57	.00
Grade	.70	.48	17.56	.00
Fall Style D	.73	.53	21.74	.00
Increase in Style C	.74	.55	8.72	.00
Fall Style G	.75	.56	4.20	.04
SES	.75	.56	3.63	.06

Coping Style score and the eleventh and subsequent dummy variable. Thus, a somewhat arbitrary limit of 5 variables was used an an "optimum" number for inclusion in the equations using Coping Styles and a 10-variable limit for the dummy variable equations.

In the dummy coded variable analyses, after Fall Reading and grade were accounted for, only a small number of significant predictors enter the equation. These are shown in Table II.

The dummy coded variables represent a pupil's predominant Fall and predominant Spring Coping Styles as *one* variable. E.g., B-H shows a pupil who changed *from* inappropriately self-directed in the Fall *to* other-directed, task oriented in the Spring.

DISCUSSION

With the major contribution of Fall Reading achievement and grade level to the variance in Spring Reading achievement recognized and accounted for above, the contribution of other variables will be discussed below.

Table II

Significant Dummy-Coded Predictors of Reading Achievement

Variable	Multiple R	Multiple RSQ	F to Enter	Prob
Teacher Directed Setting				
Fall Reading	.62	.38	213.02	.00
Grade	.64	.41	14.90	.00
SES	.64	.42	2.38	.12
C-C	.65	.42	3.49	.06
B-H	.65	.43	3.58	.06
D-D	.66	.43	3.77	.05
Program Directed Setting				
Fall Reading	.67	.45	185.57	.00
Grade	.70	.48	17.56	.00
D-B	.71	.51	9.71	.00
G-C	.72	.52	8.36	.00
H-F	.73	.54	8.01	.01
F-H	.74	.55	7.61	.01
D-D	.75	.57	6.10	.01
C-C	.76	.57	3.84	.05
F-D	.76	.58	4.10	.04

Teacher Directed Setting

In the Teacher Directed Setting sample, both Fall and Spring Other-directed, task-oriented (Style H) and Fall and Spring Compliant (Style E) behaviors emerged in two analyses as significant ($p < .05$) predictors of Spring Reading achievement.

The dummy variable analysis revealed that the pupils who remained Passive, withdrawn (Style C) in the TD sample were least likely to gain in Reading. Other types of pupils least likely to gain in Reading are those: who move from Inappropriately. Self-directed to Other-directed, task oriented (B-H); who remain Peer Dependent (B-D); and who move from Inner-directed, task oriented to Other-directed, task oriented (G-H).

Program Directed Setting

In the Program Directed Setting sample, different behaviors appeared as significant ($p < .05$) predictors. Low Peer Dependent (Style D) and high Inner-directed, task oriented (Style G) behaviors along with SES were significant predictors of the Fall styles. Of the Spring styles, low Passive, withdrawn (Style C) behavior and low Inappropriately Self-directed (Style B) behavior both appeared as significant predictors. A decrease in Style D from Fall to Spring and an increase in Style C from Fall to Spring were both also significantly related to Spring Reading achievement.

The dummy-coded variable analysis revealed in the PD sample, that those pupils who moved from Peer Dependent (Style D) to Inappropriately Self-directed (Style B) were least likely to gain in Reading. Other pupils in PD setting who are likely to gain *least* in Reading are those: who move from Inner-directed, task oriented to Passive, withdrawn (B-C); who remain Peer Dependent (D-D); and who remain Passive, withdrawn (C-C). Those in the PD setting who are likely to gain *most* in Reading are those: who move from Other-directed, task oriented to Social, productive (H-F); who reverse this pattern (F-H); and who move from Social, productive to Peer Dependent (F-D).

Other dummy variables entered the equations, but none were significant. The significant dummy variables, along with Fall Reading and grade level, accounted for 44% (TD setting) and 58% (PD setting) of the variance in Spring Reading achievement.

No additional Coping Styles or changes in Coping Styles were significant predictors of Reading achievement. The significant predictors, including Fall Reading and grade level, all accounted for about 43% (TD setting) and 56% (PD setting) of the variance in Spring Reading achievement.

CONCLUSION

These findings are in general agreement with other studies and hold implications for classroom teachers. It is important to note that differences in classroom setting will interact with student behavior and that differences in student Coping Style will result in differences in achievement. Treatment schedules are available (Spaulding, 1970) which will assist the teacher in moving students from Aggressive (Style A) to Social, productive (Style F). It may be as appropriate to address some energy to this domain as to the direct instruction in Reading.

REFERENCES

Camp, B. W., & Zimet, S. G. The relationship of teacher rating scales to behavior observations and reading achievement of first-grade children. *Journal of Special Education*, 1974, *8*, 353-359.

Cobb, J. A. Relationship of discrete classroom behaviors to fourth-grade academic achievement. *Journal of Educational Psychology*, 1972, *63*, 74-80.

Coker, H., & Lorentz, J. L. *An examination of student coping style, teacher control, and student achievement in reading*. A paper presented at the 25th Annual Meeting, National Reading Conference, St. Petersburg, Florida, 1975.

Dixon, W. J. (Ed.). *Biomedical computer programs* (3rd ed.). Berkeley: University of California Press, 1974.

Draper, N. R., & Smith, H. *Applied regression analysis*. New York: John Wiley & Sons, 1966.

Hieronymus, A. N., & Lindquist, E. F. *Iowa tests of basic skills*. Boston: Houghton Mifflin, 1971.

Iowa Tests of Basic Skills Forms 5 and 6. *Age equivalent/standard score tables*. Boston: Houghton Mifflin, 1973.

Lahaderne, H. M. Attitudinal and intellectual correlates of attention: a study of four sixth-grade classrooms. *Journal of Educational Psychology*, 1968, *59*, 320-324.

McKinney, J. D., Mason, J., Perkerson, K., & Clifford, M. Relationship between classroom behavior and academic achievement. *Journal of Educational Psychology*, 1975, *67*, 198-203.

Samuels, S. J., & Turnure, J. E. Attention and reading achievement in first-grade boys and girls. *Journal of Educational Psychology*, 1974, *66*, 29-32.

Soar, R. S. *Final report — follow through classroom process measurement and pupil growth (1970-71).* Gainesville: Institute for Development of Human Resources, University of Florida, 1973.

Spaulding, R. L. *Classroom behavior analysis and treatment.* San Jose: Author, 1970.

Warner, W. L., Meeker, M., & Eells, K. *Social class in America.* New York: Harper, 1960.

LARRY MIKULECKY
Indiana University, Bloomington, IN

ANNE WOLF
Oconomowoc Public Schools, Oconomowoc, WI

EFFECT OF UNINTERRUPTED SUSTAINED SILENT READING AND OF READING GAMES ON CHANGES IN SECONDARY STUDENT'S READING ATTITUDES

One of the dilemmas of achievement-oriented reading instruction is that such instruction can produce adult readers who are able, but seldom choose, to read. Surveys of adult reading attitudes and habits tend to indicate that the majority of adults in the United States rarely pursue reading as an independently-initiated activity. Survey information on book reading demonstrates this situation most graphically. Roos (1955) reports that only 1% of the adult population completes a book during a year, and Meade (1973) cites a Harris poll which indicates that only 26% of all adults read in a book during a 30 day period. Perhaps most startling is the Gallup Poll report (1969) that 58% of the population indicate they have "never read, never finished" a book. While the interpretation of these figures and the effectiveness of the polling questions are open to conjecture, most surveys of adult reading habits and attitudes consistently reflect a population that could be described as "aliterate," able to read, but rarely choosing to read.

There are indications that the decay or deterioration of positive attitudes toward reading evident in adult reading habits seems to be a trend that continues throughout children's school experiences. Bullen (1972) validated the *Bullen Reading Attitude Measure* with several hundred elementary students ranging from grades one through six. The mean total reading attitude scores for grades 1-3 reflect a steady drop in scores at each successive grade level. An expanded version of the Bullen measure, given to students in grades 4-6, indicated that the trend of lower mean scores at each successive grade level continued, with the sixth grade scores being the only exception. Mikulecky (1976), in norming the *Mikulecky Behavioral Reading Attitude Measure* (MBRAM) with 1,970 students from urban, suburban and rural backgrounds, noted a similar regular drop of mean reading attitude scores on the MBRAM at each successive grade level from seventh grade through twelfth grade. The trend held constant for the total sample and for each of the urban, suburban, rural sub-groups.

Secondary teachers have not been unaware of the lack of enthusiasm for reading in adolescents and adults. They have recognized that many of their students do not like to read and do not choose to read. Teachers have informally attempted to improve students' reading attitudes for quite some time. This study examines the effects upon reading attitudes of three commonly used reading-teaching activities.

Many teachers and school systems have implemented versions of Lyman Hunt's (1967) Uninterrupted Sustained Silent Reading (USSR) in an effort to

137

improve reading attitudes, reading abilities, and reading habits of students. Other teachers have incorporated the use of a growing number of commercial reading-related games as means of motivating students in reading activities and perhaps raising reading attitudes through positive game experiences. Still other teachers have had students self-select skills-related materials and activities, assuming that those students will be motivated to do the work and will therefore perform better and have a more positive attitude toward reading. The question we proposed to answer was: "What is the effect of each of these treatments upon the reading attitudes of seventh grade students?"

METHOD

Subjects

The 135 subjects were all white seventh graders attending the junior high school in Oconomowoc, Wisconsin. Subjects were students in six required quarter-credit ten-week developmental reading classes, all of which were taught by one of the researchers. Students ranged in ability from first to ninth stanines on the vocabulary and comprehension subtest of the *Stanford Diagnostic Reading Test, Level II, Form W.* No significant differences among groups were found when one way analysis of variance was performed on vocabulary and comprehension socres.

Treatments

The three treatments in the study were Uninterrupted Sustained Silent Reading, Self-Selected Reading Games, and Self-Selected Reading Skills.

Uninterrupted Sustained Silent Reading. Subjects in this treatment group simply read books of their own choosing in class for approximately 20 minutes a week. Subjects were given the choice of bringing their own books or of selecting from Scholastic Action libraries 1a-4a which were provided for them. During the quarter, book talks were given about the *Action* books. If students wished to borrow books, they were free to do so.

Games. Subjects in this treatment group played reading-related games in class for approximately 20 minutes a week. The following games were provided for subjects: *Scrabble, Ad Lib, Spill-and-Spell, Mad Libs, Password, Probe,* and *Silly Syntax.* Subjects were given a games talk at the beginning of the quarter and then selected their own games and playing partners at each session.

Skills. Subjects in this treatment group participated in self-selected skills activities in class for approximately 20 minutes a week. The following skills-related materials were provided for subjects: *Sprint, Countdown, Six-Way Paragraphs, Go, Word Power, Timed Readings, Word Mazes* and *Puzzles: 3, Feedback,* and *Brain Teasers.* At the beginning of the quarter, subjects were given a materials talk and were allowed to select their own materials. They worked alone, in pairs, or in small groups as they chose. Work was not handed in or graded.

Procedures and Measures

Treatments were administered for 20 minutes a week as a required part of a ten-week developmental reading course taken by all seventh graders at

Oconomowoc Junior High School. Two of the six classes taking the required course were randomly assigned to each of the three treatment groups. During the quarter regular classroom work was skills-oriented and focused on word analysis and comprehension.

At the beginning of the 10-week quarter, all subjects were pretested with the vocabulary and comprehension subtests of the *Stanford Diagnostic Reading Test, Level II, Form W.* This information was used to establish the fact that the three treatment groups were equivalent in reading ability. Attitude toward reading before and after treatment was measured with the *Mikulecky Behavioral Reading Attitude Measure* (MBRAM). This is a 20-item measure which reflects Krathwohl's Taxonomy of the Affective Domain, using behavioral descriptions of life-like out-of-school reading situations. For example,

13. You find yourself giving special books to friends or relatives as gifts.

VERY UNLIKE ME	1	2	3	4	5	VERY LIKE ME

The MBRAM has been validated against eight other reading attitude indicators and demonstrates a test-retest reliability of .9116 (Mikulecky, 1976). At the end of the quarter, subjects completed a questionnaire that included items on self-perceived changes in attitude toward reading. Subjects also answered open-ended questions on what they liked most and least about reading and what they would change about the class if they could.

Data Analysis

One-way analyses of variance for the three treatment groups were run on two dependent variables: gain score differences between post-and-pretests of the *Mikulecky Behavioral Reading Attitude Measure* and self-reported ratings of changes in liking for reading. Scheffé confidence intervals were used for post-hoc comparisons of groups when significant F-ratios were calculated. All significance tests were conducted at the .05 level. All open-ended question responses which referred to any of the three treatments were tabulated. Number of subject responses indicating acceptance or rejection of treatments were compared across all three treatments groups.

RESULTS

MBRAM Gain Scores

The gain scores for all treatment groups were slightly negative, indicating a slight deterioration of positive reading attitudes. The mean gain scores from smallest to largest loss were: USSR (-.275), Skills (-1.468), and Games (-1.744). An analysis of variance was performed. The F-ratio, $F(2,123) = .19$, was not significant at the $p > .05$ level.

Self-Reported Liking for Reading

The subjects rated their own increased liking for reading on a five-point scale with 1 indicating a decrease in liking, 3 indicating no change in liking, and 5 indicating an increase in liking for reading. The mean rating for the entire group ($N = 129$) was 2.721, indicating a slight decline in liking for reading over the 10 week period. Analysis of variance of rating scores among treatment groups revealed an F-ratio significant at the $p > .05$ level, $F(2,126) = 3.85$. Post-hoc Scheffé tests revealed a significant difference favoring the skills group scores over the games group scores only.

Open-Ended Items

Subjects responded to three open-ended items: what they would like to change about reading class, which classroom activity they liked best, and which classroom activity they liked least. One hundred eleven subjects responded to the item asking what should be changed in the reading class. Of the 111 responses, 44, or 40%, related to the treatments. The remainder of the suggestions dealt with such things as "no teachers" or "no tests." Only two of the 44 suggestions were negative, requesting that USSR be eliminated. In total [illegible] subjects, or 56% of those responding with treatment-related statements, requested more USSR; 10, or 22% requested more games; and 9, or 21% requested more self-selected skills work.

One hundred twelve subjects responded to an open-ended item asking which classroom activity they liked best in the reading course. Of the 112 responses, 94, or 84%, chose the reading attitude improvement treatment activity as most liked. Subjects also responded to an open-ended item asking which activity they liked least in reading class. Of the 99 responses made, 5, or 5%, mentioned a reading attitude improvement activity as least liked.

DISCUSSION AND CONCLUSIONS

Reading attitudes, as measured by the MBRAM, deteriorated slightly in all treatment groups with no statistically significant differences among mean gain scores of groups. Though not statistically significant, the USSR treatment group demonstrated the smallest loss in mean reading attitude score and the reading games treatment group demonstrated the greatest loss. The overall decay of positive reading attitudes for all groups seems to indicate that the brief treatments were not sufficient to reverse the previously noted trend of deteriorating positive reading attitudes.

The survey item measuring self-reported liking for reading also indicated an overall slight decline in liking for reading. The skills group gave evidence of a very slight increase in liking for reading on this single item rating scale. The increase in liking for reading in the skills group may be a product of short-term successes in the structured skills activities. This same carry-over would not be as easily reflected in the 20 item MBRAM which is behaviorally oriented toward out-of-school reading attitudes.

Open-ended questions requesting suggested changes in the reading course and identification of most and least liked classroom activities revealed that more students requested further incorporation of USSR into reading class than either of the other two treatments. Such requests for USSR appeared more frequently than requests for the other two treatment activities at better than a 2 to 1 ratio.

Over 80% of the subjects indicated that the reading attitude improvement treatment was the most liked activity of their reading class. This may be related to the fact that each treatment activity was characterized by self-selection within the treatment parameters. USSR subjects chose their books, gaming subjects chose their own games, and skills subjects chose their own skill activities and materials.

The lack of statistically significant differences between treatment groups may have been a product of the briefness of treatment (20 minutes per week for 10 weeks) and the fact that reading attitudes may take a great deal of time to develop and change. Further research should be conducted to determine if this is true, since quarter credit reading classes and brief treatments are not unusual or atypical in secondary schools. Research on the effectiveness of USSR ought to continue since the data gave some indication that USSR is a popular and promising technique for retarding the decay of positive reading attitudes, and it is a treatment that can be easily administered by secondary school teachers.

REFERENCES

Bullen, G. F. *The development and validation of a reading attitude measure for elementary school children.* Washington, D.C.: Office of Education, Project No. 1A045, ERIC ED 109211, 1972.

Gallup, G. *The Gallup Poll.* New York: American Institute of Public Opinion, 1969.

Hunt, L. C., Jr. Evaluation through teacher-pupil conferences. In T. C. Barrett (Ed.), *The evaluation of children's reading achievement.* Newark, DE: IRA, 1967, 111-126.

Meade, E. L. Reading: The first R — A point of view. *Reading World,* 1973, *12,* 169-80.

Mikulecky, L. *The developing, field testing, and initial norming of a secondary/adult level reading attitude measure that is behaviorally oriented and based on Krathwahl's Taxonomy of the Affective Domain.* Unpublished doctoral dissertation, University of Wisconsin-Madison, 1976.

Roos, E. Literature in the junior high. *English Journal,* 1955, *44,* 141.

TREATMENT GROUPS MATERIALS REFERENCES

Adams, M. *Word mastery with puzzles and games: Book C.* Scholastic Book Services, New York, 1975.

Goldsweig, B. *Countdown: Scope/Studyskills T.* New York: Scholastic Book Services, 1973.

Goldsweig, B. *Sprint: Scope/Speed Reading I.* New York: Scholastic Book Services, 1970.

Herber, H. *Go: Reading in the Content Areas: Level 8.* New York: Scholastic Book Services, 1975.

Higgins, J. *Brain Teasers: Interaction Program.* Boston: Houghton-Mifflin, 1973.

Pauk, W. *Six-Way Paragraphs.* Providence, R.I.: Jamestown Publishers, 1974.

Scholastic Magazine, Feedback. New York: Scholastic Books Services, 1974.

Scholastic Magazine, Word Power: Scope/Word Skills, 3. New York: Scholastic Book Services, 1974.

Spargo, E. *Timed Readings, Level I.* Providence, R.I.: Jamestown Publishers, 1975.

GAMES REFERENCES

Ad-Lib. New York: E. S. Lowe Company, 1970.

Mad-Libs. Los Angeles: Price/Stern/Sloan Publishers, 1965.

Password. Springfield, Mass.: Milton Bradley, 1963.

Probe. Salem, Mass.: Parker Brothers, 1964.

Scrabble. New York: Selchow and Righter Company, 1953.

Silly Syntax. (Interaction Programs). Boston: Houghton Mifflin, 1973.

Spill and Spell. Salem, Mass.: Parker Brothers, 1971.

CHARLES W. PETERS
Oakland Schools, Pontiac, MI.

PREDICTING READING PERFORMANCE AT THE SECONDARY LEVEL THROUGH THE UTILIZATION OF A STUDENT SELF-ASSESSMENT INVENTORY

Most general assessment procedures at the secondary level tend to be narrowly focused. As a result, decisions about reading performance in many secondary schools are made with input from limited sources. In most cases, this tends to be a general reading test such as the Iowa Silent Reading Test, the Gates-MacGinitie Reading Tests, etc. Given these circumstances, additional methods of reading assessment must be considered since most diagnostic models (Ekwall, 1976; Otto, McMenemy, & Smith, 1973; and Peters, 1977) call for the inclusion of as many variables as possible when attempting to measure reading performance.

However, the problem at the secondary level is that few educators (Holmes and Singer, 1966) have attempted to ascertain the predictive power of the variables they utilize to assess reading performance. Peters and Peters (1976) was one of the first studies to investigate the predictability of a number of variables thought to be highly related to reading ability at the secondary level (general reading performance, specific reading performance, teacher assessment, grades, and student self-assessment). They found that when the results obtained from the cognitive self-rating scale, one component of the Student Self-Assessment Inventory, and the general and specific measures of reading performance were combined, the combination increased the variance accounting for reading achievement.

Since the cognitive self-rating scale seemed to augment the screening process, the present study sought to ascertain not only the extent to which the cognitive self-rating scale discriminates between good, average, and poor readers but also whether the information derived from this instrument is distinctly different from that contributed by two additional measures of reading performance, the Gates-MacGinitie Reading Tests, Survey E, Form 1 and an attitude inventory, the second component of the Student Self-Assessment Inventory. More specifically, answers to the following questions were sought: (1) Are the cognitive self-rating scale and the attitude inventory able to discriminate between good, average, and poor readers? (2) Are the cognitive self-rating scale and the attitude inventory measuring two separate components of the reading process? (3) Is the information contributed by the cognitive self-rating scale and the attitude inventory significantly different from the information contributed by the Gates-MacGinitie? (4) Which items on the Student Self-Assessment Inventory are the best predictors of reading performance?

PROCEDURE

Subjects

The sample was comprised of 304 subjects from two suburban Detroit high schools located in a middle class urban community. The students were enrolled in regular tenth grade English classes.

In order to determine the student's level of reading performance, each subject was administered the Gates-MacGinitie Reading Tests, Survey E, Form 1. The scores from the vocabulary and comprehension components of the test were utilized to divide the subjects into three ability groups (good, average, and poor). The raw scores from the vocabulary and comprehension components of the test were converted into percentile scores. Based upon these results, the students whose scores fell between the 16th and 35th percentiles were designated poor readers; those students who fell between the 40th and 60th percentiles were designated average readers while those students who fell between the 65th and 99th percentiles were designated good readers. Students who scored between the 1st and 15th percentiles were eliminated because they were considered to be disabled readers who lacked sufficient skills to read correctly the items on the questionnaire. Those students whose scores fell between the 36th and 39th percentiles and 61st and 64th percentiles were eliminated in order to prevent a potential overlap between the poor and average and average and good categories.

Materials and Design

The Student Self-Assessment Inventory (SSAI) contained two components: 10 items which measure one's feelings about reading, the attitude inventory (e.g. I enjoy reading) and 10 items which measure one's perceptions about his or her reading ability, the cognitive self-rating scale (e.g., in comparison to other readers I would rate myself as less than average). An internal reliability coefficient of .85 for the total scale was determined by the K-R 21 formula.

Three statistical procedures were utilized to analyze the data: a correlational matrix, a factor analysis (varimax rotated), and a discriminant analysis. The correlational matrix was utilized to ascertain the relationship between the Gates-MacGinitie and the two components of the SSAI (attitude inventory and cognitive self-rating scale). The factor analysis (varimax rotated) was used to determine whether the SSAI was measuring two distinct components (affective and cognitive). The discriminant analysis procedure was employed to ascertain the predictability of each component and the degree to which each item contributed uniquely different information.

RESULTS

An analysis of the correlational matrix reveals that the correlation between the Gates-MacGinitie (GM) and the Student Self-Assessment Inventory (SSAI) is moderately low, (the attitude inventory, -.24 and the cognitive self-rating scale, -.37) indicating that the two variables are either measuring somewhat separate elements of the reading process or that the SSAI is unrelated to reading. However, the correlation between the cognitive and affective components of the

SSAI indicates that the two variables are only moderately correlated (.51).

Even though the two components of the SSAI are moderately correlated, the results obtained from the factor analysis (varimax rotated) confirm the existence of two distinguishable components. The results show that while three prominent factors emerged, only two are separately distinguishable, a cognitive factor and an attitude factor. A more detailed examination of the data reveals that factors one and two are measuring different dimensions of the same variable. Factor two measures the positive dimension of the attitude inventory (items placed in this group were all phrased in a positive manner, e.g., "Reading is a good way to spend spare time"); while factor one measures the negative dimension of the attitude inventory (items placed in this group were all phrased in a negative manner, e.g., "There is nothing to be gained from books"). While factors three and four contain items which measure one's perceptions of his or her cognitive abilities, factor four contained only one item and was considered to be insignificant.

Since the SSAI is measuring two separate factors, it had to be determined whether they both discriminated between the three ability groups. The following data is a summary of those results:

1. Student placement based upon results obtained from the cognitive self-rating scale (initial groupings based upon results from the Gates-MacGinitie).
   a. Good readers accurately placed — 91.8%
   b. Average readers accurately placed — 68.1%
   c. Poor readers accurately placed — 65.1%
2. Student placement based upon the results obtained from the attitude inventory (initial groupings based upon results from the Gates-MacGinitie).
   a. Good readers accurately placed — 80.4%
   b. Average readers accurately placed — 24.5%
   c. Poor readers accurately placed — 15.6%

In analyzing the degree to which the cognitive self-rating scale does discriminate between good, average, and poor readers, it appears that this component of the SSAI is more accurate in identifying good readers than either average or poor readers. However, when the scores derived from the attitude inventory and the cognitive self-rating scale are considered together, the ability of the instrument to discriminate between average and poor readers increases.

An item analysis revealed that six items are the best predictors of reading performance. Each item not only discriminates between good, average, and poor readers, but also provides significantly different ($p. < .001$) information from that which is contributed by the GM. In addition each of the items is from the cognitive self-rating scale which provides further support for the generalization that the cognitive self-rating scale is measuring distinctly different information from that contributed by the GM.

CONCLUSIONS AND IMPLICATIONS

As the results from this study indicate, the SSAI provides information that is distinctly different from a measure of general reading ability. Results obtained from the discriminant analysis and the correlational matrix confirm this finding. By augmenting the initial screening process, which generally includes only a measure of general reading performance with an instrument which assesses one's feelings about his or her own ability, tentative judgments about reading performance at the secondary level can be more accurate. Furthermore, it provides teachers with two distinctly different types of diagnostic information.

However, the entire instrument does not need to be utilized; only the cognitive self-rating scale since it is a better discriminator of reading performance than the attitude inventory. This finding is consistent with previous research (Dulin & Chester, 1974; Estes, 1971; Mathewson, 1976; and Peters & Peters, 1976) which has shown that one's attitude toward reading is unrelated to one's reading ability. These results contribute further support for that generalization.

This study demonstrates that a cognitive self-rating scale can have an additional function not previously considered by many researchers or classroom teachers, but the findings are by no means conclusive. Indeed at least two questions generated by this study remain unanswered: (1) Why do poor readers appear to be more "honest" in their responses to self-assessment questions that measure cognitive abilities in comparison to those questions which seek to assess their feelings about reading? In this study, for example, poor readers tend to answer the questions from the affective component the way they believed the teacher wanted them to answer the questions. For example, 63 percent of the poor readers felt that reading was a good way to spend their spare time. Good readers did not even respond that positively to this particular statement thus suggesting that poor readers possibly believed that this was a more "appropriate" response. (2) Can the SSAI replace instruments such as the GM in testing situations that require only initial decisions about general reading performance, i.e., whether a person appears to be a poor reader and, therefore, needs more intensive diagnosis?

Further research needs to be conducted to ascertain whether a 10 item cognitive self-rating scale might be just as reliable as the GM in placing students according to general reading ability. The implications are significant for classroom teachers; because rather than spending 50 minutes administering a general reading test such as the GM, they might be able to use a cognitive self-rating scale which would not only require less time to administer but would also provide results which are comparable to those obtained from a general measure of reading performance.

REFERENCES

Dulin, K. L. & Chester, R. D. A validation study of the Estes Attitude Scale, *Journal of Reading*, Vol. 15, No. 1, (October, 1974), 56-59.

Ekwall, E. *Diagnosis and remediation of the disabled reader.* Boston: Allyn and Bacon, 1976.

Estes, T. H. A scale to measure attitude toward reading. *Journal of Reading*, Vol. 15, No. 2, (November, 1971), pp. 135-38.

Holmes, J. & Singer, H. Speed and power of reading at the high school level, *U.S. Office of Education Monograph Series*, 1966.

Mathewson, G. C. The function of attitude in the reading process. In *Theoretical Models and Processes of Reading*, 2nd Ed., H. Singer and R. Tuddell, (Eds.), Newark, Delaware; International Reading Association, 1976, 655-676.

Otto, W.; McMenemy, R.; & Smith, R. *Corrective and remedial teaching*, (2nd Ed.), Boston: Houghton Mifflin Co., 1973.

Peters, C., & Peters, N. A systematic approach to predicting reading performance at the secondary level. In W. D. Miller & G. H. McNinch (Eds.), Reflections and Investigations on Reading. *Twenty-fifth Yearbook of the National Reading Conference*, 1976, 121-128.

Peters, C. Diagnosis of reading problems. In W. Otto, N. Peters, and C. Peters (Eds.), *Reading problems: A multi-disciplinary perspective*. Reading, Massachusetts: Addison-Wesley, 1977, 151-188.

PATRICIA L. ANDERS
University of Arizona

KEN L. DULIN
University of Wisconsin, Madison

MASCULINE, FEMININE, AND ANDROGYNOUS COLLEGE-AGE READERS' STORY-PLOT PREFERENCES RELATED TO GENDER OF PROTAGONIST

The existence of sex-related patterns of preference in reading materials, with males preferring violent, action-packed stories of war, crime, and one-to-one competition, and females favoring optimistic, romantic, "and they lived happily ever after" domestic-dramas, has been reported over and over in the professional literature of reading. "Boys will be boys," the conclusion has been drawn, "and girls will be girls, even in their reading activities."

Of late, however, this traditional view has come to be questioned, with the notion being advanced that perhaps *psychological* patterns, not simply physical-gender ones, are responsible for patterns of behavior similar to these. One artifact of this questioning has been the Bem Scale for measuring sex-role identity.

The Bem Scale, developed by Sandra Bem, a Stanford University psychologist (Bem, 1974), purports to measure "androgyny," a condition Bem postulates as that of an individual who is psychologically *neither* traditionally "masculine" nor traditionally "feminine," but rather tends to incorporate the traits associated with *both* of those concepts into his or her personality.

To take the Bem Scale, a subject rates 60 "traits" on a seven-point scale running from a 1 of "never or almost never true of me" to a 7 of "always or almost always true of me," after which these scores are summed and compared to give measures of "masculinity," "femininity," and "androgyny," with individuals falling into this final category by being above the group mean in both "masculinity" and "femininity."

METHOD

This study attempted to assess college-age readers' sex-of-protagonist preferences in story-plots as these preferences were related to their Bem Scale placements. Ten book-annotations previously validated as being primarily either "masculine" or primarily "feminine" in their appeal were re-written in two forms each, with the five "masculine" ones and the five "feminine" ones thus each occurring with both male and female protagonists. The annotations were then offered in alternate forms of a questionnaire to 84 subjects who rated them on a five-point scale in terms of how much they would like to read them.

In addition, these subjects were all given the Bem Scale, and the results of the two data-collections were compiled and statistically analyzed. The method of analysis in all cases was One Way Analysis of Variance. Given the descriptive nature of the study a .05 level of significance was deemed appropriate for purposes of interpretation.

RESULTS

First, the total group of subjects', the male subjects', and the female subjects' differences in responses to each of the five male-interest-oriented book titles and annotations and to each of the five female-interest-oriented book titles and annotations were compared. Only one of the thirty comparisons resulted in a statistically significant difference. Thus, it can be quite safely concluded that the differences between the total group responses, the male subjects' responses, and the female subjects' responses to each of the book titles and annotations were not significant.

Second, the androgynous males and females and the non-androgynous males and females differences in responses to each of the masculine-interest-oriented book titles and annotations and the female-interest-oriented book titles and annotations were compared. Again, only one of the thirty comparisons resulted in a statistically significant difference. Thus, at the group level at any rate, it can be concluded that sex-role identity (androgynous or non-androgynous) has little to do with preference for masculine-interest-oriented books or for feminine-interest-oriented books.

Next, male subjects and female subjects pooled ratings of the five male-protagonist and five female-protagonist versions of the ten simulated book annotations were compared (summary statistics are shown in Table 1). Here, the results

Table 1
Summary Statistics of the Male and Female
Subjects' Pooled Raw Scores Across
Passage Form and Version

Form/Version	N	Mean	Standard Deviation
One/Male[a]			
Male	17	12.47	3.96
Female	24	14.25	3.58
One/Female[b]			
Male	17	12.35	4.17
Female	24	17.92	3.06
Two/Male[c]			
Male	16	13.81	4.12
Female	27	16.67	3.52
Two/Female[d]			
Male	16	14.00	4.47
Female	27	13.11	3.61

a Stereotypic "masculine" annotations with male protagonists
b Stereotypic "feminine" annotations with female protagonists
c Stereotypic "feminine" annotations with male protagonists
d Stereotypic "masculine" annotations with female protagonists

are more interesting. Two *F*-ratios were found to be significantly different, (*F* (1, 39) = 24.33 and *F* (1, 42) = 5.82): the female subjects preferred the five titles and annotations that conformed to the stereotypic "feminine" interests, regardless of male or female story-protagonist. Two *F* ratios were *not* found to be significant: no statistically significant differences were found between subjects when subjects judged stereotypic "masculine" interests. Thus, the gender of the protagonist made no difference to the subjects; rather, it was the content of the annotations that made a difference. Specifically, female subjects responded differently than did the male subjects of stereotypically "feminine" annotation-content regardless of the sex of the protagonist.

Finally, the "androgynous" and "non-androgynous" subjects' pooled ratings of the five male-protagonist and five female-protagonist versions of the ten simulated book-annotations were compared. As was found in the previous sex-role-identity comparisons, no significant differences occurred.

CONCLUSIONS

Overall, then, the usefulness of the Bem Scale as a predictor of reading interests seems definitely to be in doubt. Subject sub-groupings obtained seem not to vary much in terms of their response to the characteristics of story-materials, at least in terms of the sex of the protagonist.

Physical gender, however, or simply "sex" of the subject in the traditional sense of the term, did seem to make a difference; this distinction — and it alone — gave us our only significant differences in pattern of response, with the *females* responding in the most "traditional" manner. That is, the females preferred romantic and domestic-type book annotations whether the protagonist was a male or female.

IMPLICATIONS

Where, then, from here? Not the Bem Scale, apparently, even though much of the earlier literature on it had seemed to be so very promising. Perhaps a *new* scale, built less on the patterns of the past, will emerge, or maybe we will just have to settle for that *older* set of distinctions (if distinctions must be made): male and female. Perhaps they *are* the real "difference," "vive" it or not.

REFERENCE

Bem, S. L. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 1974, *42*, 155-162.

ANTHONY V. MANZO
University of Missouri, Kansas City

A PRELIMINARY STUDY OF THE RELATIONSHIPS AMONG ANIMAL IDENTIFICATIONS, LEARNING PREFERENCES, READING, AND READING RELATED ABILITIES

Always there is the nagging issue of who and what a student is other than his achievement scores. No one argues any longer that the shape of "success" and "failure" have been too easily defined in terms of academics; a goodly number of the high achieving are social-emotional losses, and an equal number of the median achieving, or as Harry S. Truman referred to them and himself as "straight C students," assume the reins of leadership. Arguments over whether or why this should be the case are worthwhile, but as the arguments continue, so too, does the prevailing dynamic. It was not as an answer to this difficult question, but as a pragmatic means of getting a purchase on it, that the current line of research was undertaken. Sought is a means of identifying some of the social-psychological characteristics related to academic success and lack of such success. The continuation of this research will attempt to identify such other aspects of motivation and style as may attenuate "success" in *non-academic* affairs. Two experimental instruments have become central to this line of research, one a measure of social-psychological characteristics, the other an inventory of styles of teaching-learning.

This paper reports the partial results of a correlational study of the interrelationships of these instruments with reading, and reading related academic skills. Several specific and consistent, though relatively weak, patterns of relationship were found.

METHOD

Procedure

The data of this report were derived from two sub-samples of ninety-four (94) and thirty six (36) cases of 10th grade students, drawn from over one thousand, randomly selected students tested at grades 7 thru 12 at a large midwestern, predominately white, school district.

Assessed in this correlational study were the following: 1) general social-psychological characteristics as these could be inferred from selected animal identifications on the Manzo Bestiary Inventory (MBI), 2) teaching-learning preferences as measured by the Learning Preference Inventory (Manzo, 1976 revision), 3) cloze passage performance on two passages, one at grade level and another at a more advance level drawn from a popular social studies textbook, 4) multiple choice comprehension and vocabulary in context scores on these same two passages, 5) writing abilities (subjective assessment by two judges based on selected criteria), and 6) stated intentions to either drop-out of high school, complete high school, do technical training, complete college, or do post-graduate study. Sex differences were also noted.

Instruments

The Manzo Bestiary Inventory (MBI), group form, is a social-psychological inventory in which subjects express degrees of identification with thirty-three different animals. Expressions are recorded on a five point scale, one low, five high. Personality, social, and behavioral characteristics are inferred from animal identifications in essentially two ways: by the face values and corresponding metaphoric meanings typical to animal imagery, and by a more elaborate, collateral method by which a rank ordering of previously collected associations, or 'adjectivals', is examined for the meanings these might suggest.

The Learning Preference Inventory (LPI) is the second experimental instrument. The LPI asks students first to indicate their degree of "Knowledge and Experience" with, and then their preferences for, ten fundamental methods of teaching-learning. The ten methods are repeated for several school subjects to determine the extent to which preferences vary by subject/task requirements. The latter findings are not included in this report. The ten general methods of teaching-learning offered for choice on the inventory are expressed below.

1. *Lecture Learning:* the teacher does most of the talking. Questions are permitted, but there is generally little discussion.
2. *Lecture/Discussion:* the teacher lectures briefly, then raises questions for class discussion.
3. *Inquiry Lesson:* the teacher briefly describes a topic; the class questions the teacher in an effort to discover the important information. The teacher finally tells the class what important questions they may have failed to ask.
4. *Incidental Teaching:* most skills and some information are taught as they appear to be needed. For example, a spelling, reading, or writing lesson may be taught on the spot in a social studies class because the students seemed to need it.
5. *Casual Learning:* specially designed games are played or activities done to improve certain skills, encourage certain attitudes, or build knowledge in certain areas.
6. *Directed Individual Learning:* a student is given certain work to do based on tests which showed his special needs. He must complete this work before he can go on to a higher level.
7. *Student Reporting:* individual students are responsible for finding and presenting information to the class on assigned topics.
8. *Group Work:* small groups are assigned topics to research and/or discuss. Their findings and/or conclusions then are shared with the entire class.
9. *Individual Tutoring:* the teacher works with one student (or a very small group) while the remainder of.the class is engaged in another activity.
10. *Team Teaching:* two or more teachers working together teach a class. They discuss different points of view in front of the group. The class may participate in the discussion.

Validity and Reliability of Experimental Instruments

Full information regarding the validity and reliability of the MBI & LPI are still pending. However, the reasonability of employing these instruments, based on logic and partial data, can be stated. One reason for employing the instruments is the fact that this study itself is a means of collecting validity and reliability data.

The justification for the LPI is fashioned from the fact that there are relatively few different categorical methods by which to teach and learn in school. By stating these simply, it is assumed to be possible to discover student knowledge of, and preference for these. Preferences have tended to appear relatively stable on a test-retest basis. Reliability estimates on one occasion, with 65 high school youngsters, were as high as .67, and on another with approximately 100 college students, as low as .48. An inspection of the data showed the shifting in the general preference items (those used in this study) to be concentrated to within one point 88 per cent of the time.

The concurrent validity of the LPI is difficult to establish because it is not comparable to any other instrument in its purpose. A study, in progress, does relate it to two "learning style" inventories, and to several academic measures, which should also reveal any predictive value the instrument may have along those lines.

The MBI draws part of its construct, and in a sense, content, validity from two sources: the wide scale use of animal metaphors in literature and folklore, and a study by the author which verified a very high degree of agreement among a sampling of over two hundred (200) people as to the free associations they had with the animal labels used in the instrument (Manzo, 1975).

Concurrent validity of the MBI was partially established on a sample of eighty-three (83) adults enrolled in basic education classes. Protocols from the MBI were found to be comparable to those from the *Luscher Color Test* (Scott, 1969), a projective test developed in Germany and extensively used in a variety of ways for over twenty years (Manzo, Lorton, Condon, 1975). A fuller measure of concurrent validity is pending findings from the studies, noted above, now in process. These studies relate the MBI (& LPI) to selected personality traits, as measured by such traditional measures as the *Ominbus Personality Inventory*. Findings should help to pin-point, in more familiar terms, the subtle human traits which may be related to specific animal identifications and learning preferences.

The predictive validity of the MBI will take some time to complete. This study attempts to establish the predictive validity of the MBI for matters pertaining to reading and academic success. The hope of eventually being able to predict who will be "successful" outside of school, and who might do considerably better in school if their different styles of being were met, is the extensive research project of which this report is only a small part.

The reliability and likely error variance of the MBI is calculated for each testing situation by correlating three test items which are repeated in different locations on the test. For this sample the repeated items correlated .67, .74, and .88. On the bases of these correlations, test reliability for this assessment situation was estimated at about .76.

RESULTS

The enormity of the correlation matrix plus the complexities involved in representing trends which were either reinforced or diluted by cross-verifying factors suggested this practical resolve. The data are compressed into matrices, only slightly interpolated by the experimenter, to form a display, which it is hoped, provides a better view of the subtleties disclosed than would a traditional table of statistical findings.

Matrix 1 displays the relationships (correlations) found between each MBI choice with each of the major academic factors assessed. Where associations were judged to be relatively highly related (with significance ratios beyond the .02 level, and/or cross-verified by an additional corresponding measure) it is shown as a plus (+) or minus (-) three (3) relationship. Where a relationship was found to be significant at the .03 to .05 levels, and/or showed a strong corresponding trend among interrelated items, it is shown as a plus (+) or minus (-) two (2). Where a correlation suggested a relationship which was "non-statistically significant," though a solid trend (.06 -.08), with corroborating evidence, it is noted as plus (+) or minus (-) one (1). Blank spaces indicate no apparent relationship.

Matrix 2, constructed in the same fashion, depicts relationships between the MBI and major learning preferences.

Matrix 1

Relationships Among Animal Identifications and Academic Factors[a]

	Sex Related	Cloze	Writing	Comprehension	Vocabulary	Intentions
1. Horse	.15 (F) (1)	.25 (2)		.24 (2)		
2. Dog		.24 (2)				.24 (2)
3. Squirrel		-.28 (-2)				
4. Pheasant		-.17 (-1)	-.20 (-2)		.19 (1)	
5. Cow				-.20 (-1)		
6. Rooster						
7. Lion			-.32 (-2)			
8. Chicken		.39 (3)				
9. Moose	-.15 (M) (1)		-.22 (-1)	.21 (1)		
10. Penguin	.15 (F) (1)	.26 (3)				.21 (1)
11. Owl		-.15 (1)				
12. Badger	-.22 (M) (3)	-.18 (-1)	-.24 (-1)			
13. Tiger			-.20 (-1)	.24 (2)	.26 (2)	
14. Goat						
15. Turtle		.25 (2)				
16. Alligator	-.22 (M) (3)					
17. Porpoise		.23 (2)		.23 (2)	.24 (2)	.19 (1)
18. Leopard		.25 (2)	-.24 (-2)			.32 (3)
19. Hog				-.29 (-2)		
20. Wolf	-.16 (M) (1)	-.40 (-3)	-.35 (-3)			
21. Giraffe	.25 (F) (3)	.34 (3)		-.20 (-1)	.20 (1)	.19 (1)
22. Mink	.17 (F) (1)	.19 (1)		-.22 (-1)		
23. Duck			-.21 (-2)		-.19 (-1)	
24. Peacock	.43 (F) (3)	-.17 (-1)				
25. Swan	.45 (F) (3)	.18 (1)	-.20 (-1)			
26. Falcon	-.22 (M) (3)	-.19 (-1)	-.20 (-1)		.18 (1)	.17 (1)
27. Dove	.52 (F) (3)	.30 (3)				
28. Snake	-.24 (M) (3)		-.19 (-1)	-.18 (-1)		
29. Elephant		.19 (1)		.19 (1)	.18 (1)	.18 (1)
30. Eagle	-.16 (M)(1)			.22 (2)	.24 (2)	
31. Hippopotamus		.19 (1)	.19 (1)			
32. Fox			-.19 (-1)	-.25 (-2)	.26 (-2)	
33. Coyote		-.22 (-2)	-.33 (-3)			

[a] $N = 36$. F = female, M = male, 1 = weak, 2 = moderate, 3 = strong relationships

Matrix 2

Relationships among Animal Identifications and Learning Styles[a]

	1 LL	2 Lec/Dis	3 Inq	4 Incid	5 Casual	6 Indep	7 % Reptg	8 Group	9 Tutorg	10 Team
1 Horse		20 (2)	20 (2)	18 (2)			16 (1)		.16 (1)	.19 (-2)
2 Dog		26 (3)						18 (2)		
3 Squirrel				.19 (2)					-.18 (-2)	-.18 (-2)
4 Pheasant										
5 Cow										
6 Rooster										
7 Lion								-.19 (-2)		
8 Chicken		-.18 (-2)					17 (2)			
9 Moose		.15 (1)			21 (3)					
10 Penguin					29 (3)			.25 (3)		
11 Owl					16 (1)			.16 (1)	.16 (1)	
12 Badger	-.19 (-2)		-.15 (-1)		.22 (3)	.18 (2)	.24 (-3)			
13 Tiger	-.20 (-3)									.18 (2)
14 Goat										
15 Turtle	-.22 (-3)					16 (1)				-.18 (-2)
16 Alligator										
17 Porpoise		22 (3)			21 (3)	18 (2)				
18 Leopard				21 (3)	.31 (3)	.17 (1)	-.19 (-2)			
19 Hog			17 (-2)				.18 (2)			.19 (2)
20 Wolf				16 (1)		16 (1)				
21 Giraffe		18 (2)	.20 (3)		.18 (2)			.16 (1)		-.18 (-2)
22 Mink	-.20 (-2)									
23 Duck					21 (3)				.16 (1)	
24 Peacock					24 (3)					
25 Swan					24 (3)					
26 Falcon		17 (1)								
27 Dove										
28 Snake				16 (1)						
29 Elephant										
30 Eagle										.16 (1)
31 Hippopotamus							.19 (2)			
32 Fox										
33 Coyote		17 (1)	.22 (3)							
34 Cloze	17 (-1)									
35 Reading (Voc. & Rdg.)	19 (2)	20 (3)		-.16 (-1)	19 (2)					
36 Sex		21 (M) (3)			.19 (M) (2)					

[a] N = 94. 1 = weak, 2 = moderate, 3 = strong relationship

The strongest relationships between MBI choices and academic skills are here itemized:

1. Those identifying with Horse and Porpoise have the highest and most consistent patterns of achievement. Those identifying with Eagle ran a close third.
2. There are affective characteristics associated with Horse, Chicken, Penguin, Giraffe, and to a slightly lesser degree, Dog, Turtle, Porpoise and Dove which appear to be positively related to cloze passage performance.
3. Conversely, there are affective characteristics associated with Wolf and Coyote which appear to be negatively correlated with cloze passage performance.
4. Those identifying with Fox appear to have poor writing, vocabulary and comprehension skills, while those identifying with Hog merely have poor comprehension skills.
5. Those identifying with Eagle, Porpoise, Tiger, and Horse tend to have good comprehension skills, while each of these, excepting Horse, have above average vocabulary skills as well.

The reader is referred to Matrix one to study sex differences in choices and other consistent, though non-statistically, significant trends.

Relationships of animal choices to learning preferences are here summarized:

1. *Lecture Learning* was held in highest regard by those identifying with Rooster, and lowest regard by those identifying with Tiger, Turtle, Mink, and to a lesser degree, Badger. Those identifying with Lecture Learning had above average levels of performance on reading measures.
2. *Lecture Discussion* was held in highest regard by Horse, Dog and Porpoise, to a somewhat lesser degree by Giraffe, and negatively by those identifying with Chicken. Those preferring Lecture Discussion had the highest levels of reading performance of all groups tested.
3. *Inquiry Lessons* were held in high regard by Giraffe, and to a lesser degree, Horse. Those identifying with Hog disliked this teaching-learning style.
4. *Incidental Learning* was held in highest regard by Leopard and Coyote, in high regard by Horse, Squirrel and Pheasant, and enjoyed no evidence of rejection by any discernable grouping of students.
5. *Casual Learning* was held in the highest regard by those identifying with Moose, Penguin, Badger, Porpoise, Leopard, Duck, Peacock, and Swan, and in high regard by those identifying with Owl, Tiger and Giraffe. Casual Learning enjoyed no sharp expressions of rejection. Those preferring Casual Learning tended to be above average achievers on the reading measures.
6. *Independent*, or *Directed Individual Learning*, was preferred by those identifying with Badger and Porpoise, with no group expressing rejection.
7. *Student Reporting* was most popular with those identifying with Cow, and only slightly less popular with Chicken and Hog, (and surprisingly), Fox. It was most disliked by Leopard.
8. *Group Work* was very popular with Penguin, slightly less so with Squirrel, and disliked by Rooster.
9. *Individual Tutoring* was liked by those identifying with Dove and disliked by those identifying with Pheasant.
10. *Team Teaching* was preferred by Badger and Hog, and disliked by Horse, Squirrel, and Giraffe.

The reader is referred to Matrix 2 for trends and sex differences.

DISCUSSION

There is a slight, though consistent triangulation among social-psychological characteristics, learning style preferences, and various aspects of reading-language skills. Though the actual correlations are relatively low, the relationships that are revealed are quite precise and consistent with the clinical impressions gathered over recent years with the experimental instruments. There is, for example, clinical consistency with the findings of positive correlations between the identification Eagle with comprehension and vocabulary, however, the lack of correlation between Eagle and cloze performance and Eagle and writing ability is not consistent. The negative correlations of the identification

Fox with comprehension and vocabulary, along with a negative trend in writing, and no correlation with cloze performance is most consistent with clinical findings. Consistent too, are the findings of Dove, as positively correlated with cloze performance, but with no other measured skill, and Giraffe as highly positively correlated with cloze, but somewhat negatively correlated with comprehension. Thus, the slight, though quite distinctive patterns of academic development revealed in these animal identifications appear to be sound.

It should be noted that an actual MBI test protocol is considerably easier to interpret than would be suspected from the data support levels shown for each single identification. The opportunity to study patterns of choices permits the examiner to cross-verify and temper impressions. The balance struck, for example, among the high positive choices *porpoise, eagle, fox* and *dove* is almost always indicative of a person with sound academic skills and skill in the long-term management of people. Such a person is likely to be capable and ambitious, but not driven. A similar profile in which the *dove* is replaced by a choice such as *alligator* has tended to reveal a more fitful pattern of academic successes and failures, with strong leadership tendencies, but a lack of durability in the management of people. The latter profile discloses a strand of aversively motivated behavior which must occasionally exhibit itself in hostile and/or self-defeating ways.

Learning preferences also seem to cluster in distinctive patterns with respect to both achievement and bestiary identifications. The profile of character traits and patterns of achievement, for example, of those electing 'Lecture Learning' is quite different from those preferring 'Casual Learning'. This says that the LPI, which was intended initially as a simple inventory of learning preferences, is clearly revealing of much deeper characterological traits. This could have significant implications for "learning styles" research. These findings, along with the innocuous nature of the instruments, and their ease of administration, suggest that we may be able to discover "who and what" the particular students are who are helped, or possibly inhibited in their learning, by our various treatment conditions.

A study of the learning preferences of Adult Basic Education (ABE) students mentioned earlier bears witness to the latter supposition. The life experiences, personality traits, and expressed learning preferences of the ABE students studied were all consistent. They preferred to learn by two of our most directive and efficient means, *lecture* and *tutoring*. An inspection of ABE programs operating around the country reveals, however, that students are deprived of this entry level need and expectancy. They are left to languish in a sea of "individually prescribed, independent study" materials. The idea may not be incorrect in itself, but the match is all wrong. Surely, at least a partial answer to the question of why we cannot keep adults in reading programs is nested in this fact. There likely are many other such fundamental facts which are escaping our understanding for lack of instrumentation to reveal them; which again points to the purpose of the current line of inquiry and research, to develop instruments and further knowledge of some of the more subtle, though potentially significant needs of our students.

---

*Experimental forms of the MBI and LPI are available at no charge upon request from the author.

REFERENCES

Manzo, A. V., Lorton, M., & Condon, M. Personality Characteristics and Learning Style Preferences of Adult Basic Education Students. *Research Monograph, Center for Resource Development in Adult Education,* University of Missouri, Kansas City, 1975.

Manzo, A. V., Manzo Bestiary Inventory. *Monograph of Center for Resource Development in Adult Education,* University of Missouri, Kansas City, 1975.

Scott, I., *The Luscher Color Test.* New York: Random House, 1969.

ALBERT J. KINGSTON
TERRY L. LOVELACE
MOLLY M. WILSON
The University of Georgia

ISSUES OF SEXISM AND ETHNICITY IN READING

As judged from the volume of published papers in the professional literature, there is considerable interest in the degree to which basal readers and school textbooks reflect discriminatory ethnic and sexist attitudes. Kingston and Lovelace (1977a), who recently described the methods employed in studying sexism as reflected in 78 published papers, found a variety of techniques employed to determine sex discrimination. Among the more common techniques were: a) counting the number of male and female characters (Committee on Women in Words and Images, 1972; Schnell and Sweeney, 1975), b) analyzing the differences in the ages, occupations, and actions of males and females (Frasher and Walker, 1972; Frisof, 1969; Hillman, 1974), c) enumerating the number of "male" and "female" pronouns (Gershuny, 1975; Rhome, 1971), d) determining the proportions of male to female authors (Hurst, 1973; Nilsen, 1972), or e) searching for examples where women were depicted in traditional, social or "stereotyped" roles, e.g., mothers (Weitzman, Eiffer, Hokada, and Ross, 1972; Prida, Ribner, Davilla, Garcia, Pingdollers, and Rivera, 1976). Although all types of educational materials have been examined by those pursuing this type of study, basal readers, English and social studies texts have been special targets. The present writers recently completed an examination of 123 articles on ethnicity to ascertain the methodology used in those studies, the criteria employed, and the results obtained in order to compare such works with those by authors searching for sex discrimination. (Unfortunately, limitations of space and time prevent a full discussion of these findings, so only a few of the more typical reports are cited here*).

In the case of research on ethnicism, few studies of an empirical nature were found. The majority of authors failed to present data supporting their findings or to provide evidence to support the reliability and validity of either their methods or results. Like those interested in discovering evidence of sex discrimination, these authors also: a) counted the number of minority characters depicted in books in contrast to the number of white, middle-class characters (Britton, 1974; Gast, 1967), b) examined the occupational and social roles of ethnic groups in both narrative content and illustrations (Jones, 1971; Allen, 1971), and c) searched for examples where ethnic minority characters were presented unrealistically (Katz, 1973; Hata and Hata, 1974; Falkenhagen, Johnson, and

---

* Those desiring a copy of the complete bibliography should send a large, stamped, self-addressed envelope to the senior author.

Balasa, 1973). In fact, if one were to accept the criticism of certain advocates, textbook writers are "damned if they do and damned if they don't." Some critics argue that an insufficient number of minority characters portrayed in a story represents obvious discrimination, while also arguing that white authors are incapable of depicting the true character and lives of minorities (MacCann, 1972; Arnez, 1969). Although all ethnic groups have been studied, the majority of articles deals with treatment of blacks. Blacks, of course, represent the largest minority group. Often the researchers combine concern for blacks with studies of feminine or sexist concerns (Prida et al., 1972; Racism and Sexism Resource Center for Educators, 1976). One promising fact is that a number of writers note that modern authors and publishers seem to be somewhat fairer in their treatment of ethnic minorities than in previous years. Still, they believe that both representation and treatment can be further improved (Larrick, 1975; Baronberg, 1971; Grambs, 1972; Kraus, 1975).

Because of the sheer volume of articles dealing with alleged sex and ethnic bias, most publishers and writers are cautious concerning this sensitive area. Few writers have discussed the implications of control suggested by the more militant ethnic and feminist spokesmen. Kingston and Lovelace (1977b) have raised the question of censorship. Latimer (1976), in contrast, argues that offensive books must be removed from classrooms and library shelves and that this act is not censorship but "selectivity." The implication is that such actions are necessary and beneficial. The ethics and problems involved in choosing books and materials for diverse populations undoubtedly need more discussion by educators and laymen. How can a proper balance be achieved? Even now a few educators are concerned with the fact that overzealous advocates may cause problems of reverse bias (Elkind, 1972). Materials that violate the moral, religious, and ethnic values of students have in the past caused problems in schools. They undoubtedly will in the future.

Similarly, most educators probably would reject any instructional materials selected primarily because of their propaganda value. Yet some writers seem to come close to advocating such criteria. Finally, educators long have recognized that there is a limit to which books used in the classroom can depict realistically all aspects of a large and varied nation. Sex, violence, drugs, alcoholism, crime and poverty are unfortunately aspects of modern life. However, when and how children are introduced to them constitutes a major problem in itself. Few psychologists, sociologists, or educators would advocate the use of such themes in basal readers or literature selections.

Certainly the alleviation of sex and ethnic biases is a worthy goal. Unfortunately the critics have held forth at length, with solutions ranging from censoring alleged sexist and racist works to forcing writers to adhere to certain standards. Almost all suggest quotas and idealized roles and personality characteristics for women and minority group members. There have been few rebuttals or alternate suggestions. Academia and its forums for debate have been strangely quiet. Surely this important topic needs rational discussion before much of the literary tradition and knowledge of what motivates and interests youth is destroyed.

REFERENCES

Allen, V. S. An analysis of textbooks relative to the treatment of black Americans. *Journal of Negro Education,* 1971, *40,* 140-45.

Arnez, N. L. Racial understanding through literature. *English Journal,* 1969, *58,* 56-61.

Baronberg, J. Black representation in children's books. *ERIC-IRCD Urban Disadvantaged Series #21,* May 1971 (ERIC Document Reproduction Service No. ED 050 188).

Britton, G. E. Danger: State adopted texts may be hazardous to our future. July 1974. (ERIC Document Reproduction Service No. ED 096 611).

Committee on Women in Words and Images. *Dick and Jane as victims: Sex stereotyping in children's readers.* Author, 1972.

Elkind, D. Ethnicity and reading: Three avoidable dangers. In H. Tanzer (Ed.), *Reading, children's books and our pluralistic society.* Newark, DE: International Reading Association, 1972.

Falkenhagen, M., Johnson, C., & Balasa, M. A. The treatment of native Americans in recent children's literature. *Integrated Education,* 1973, *11,* 58-59.

Frasher, R. & Walker, A. Sex roles in early reading textbooks. *The Reading Teacher,* 1972, *25,* 741-49.

Frisof, J. K. Textbooks and channeling. *Women,* 1969, *1,* 26-28.

Gast, D. K. Minority Americans in children's literature. *Elementary English,* 1967, *44,* 12-23.

Gershuny, H. L. Public doublespeak: The dictionary. *College English,* 1975, *36,* 938-42.

Grambs, J. D. & Carr, J. C. (Eds). *Black image: Education copes with color.* Dubuque, IA: 1972.

Hata, D. I. & Hata, N. I. *I wonder where the yellow went? Distortions and omissions of Asian Americans in California education.* Paper presented at the meeting of the American Educational Research Association, New York, April 1974.

Hillman, J. S. An analysis of male and female roles in two periods of children's literature. *Journal of Educational Research,* 1974, *68,* 84-88.

Hurst, G. *Sex bias in junior high school literature anthologies.* St. Louis, MO: NOW, 1973.

Jones, J. P. Negro stereotypes in children's literature: The case of Nancy Drew. *Journal of Negro Education,* 1971, *40,* 121-25.

Katz, W. A. Minorities in American history textbooks. *Equal Opportunity Review,* June 1973. (ERIC Document Reproduction Services No. ED 085 431).

Kingston, A. J. & Lovelace, T. L. Sexism and reading: A critical review of the literature. *Reading Research Quarterly,* 1977a, in press.

Kingston, A. J. & Lovelace, T. L. Guidelines for authors: A new form of censorship? *Journal of Reading Behavior,* 1977, *9,* 89-93.

Kraus, W. K. *From Steppin Stebbin to soul brothers: Racial strife in adolescent fiction.* Paper presented at the meeting of the Conference of English Education, Chicago, March 1975. (ERIC Document Reproduction Services No. ED 103 902).

Latimer, B. I. Telegraphing messages to children about minorities. *The Reading Teacher,* 1976, *30,* 151-56.

Larrick, N. *Ten years later.* Paper presented at the meeting of the American Library Association, New York, July 1975. (ERIC Document Reproduction Services No. ED 112 833).

MacCann, D. & Woodard, G. *The black American in books for children: Readings in racism.* Metucken, NJ: Author, 1972.

Nilsen, E. P. Women in children's literature. *College English,* 1972, *32,* 918-26.

Prida, I., Ribner, S., Davilla, E., Garcia, I., Pingdollers, C., & Rivera, A. Feminists look at the 100 books: The portrayal of women in children's books on Puerto Rican themes. *Interracial Books for Children,* Spring 1972.

Racism and Sexism Resource Center for Educators. *Sexism and racism in popular basal readers, 1964-76.* Washington, D. C.: Author, 1976. (ERIC Document Reproduction Services No. ED 123 307).

Rhome, F. D. *Manglish: What's it all about?* Paper presented at the meeting of the Midwest Modern Language Association, St. Louis, October 1971.

Schnell, T. R. & Sweeney, J. Sex role bias in basal readers. *Elementary English,* 1975, *52,* 737-42.

Weitzman, L. J., Eiffer, D., Hokada, E., & Ross, C. Sex-role socialization in picture books for preschool children. *American Journal of Sociology,* 1972, *77,* 1125-50.

JULES C. ABRAMS
Hahnemann Medical College & Hospital

THE MYTHS AND REALITIES OF SEVERE READING AND RELATED LEARNING DISABILITIES

The question may be raised why a conference concerned with the teaching of reading should address itself to the area of learning disabilities. On the other hand, such a question could be viewed as naive and at best irrelevant. If we define reading as "the reconstruction of the facts behind the symbols," we speak of a process that extends far beyond the mere decoding of the written word. Comprehension, organization, utilization of thinking skills, etc., are as much a part of "reading" as of "learning." If "reading is learning," then a comprehensive knowledge of the factors involved in the diagnosis and treatment of learning disorders would be tremendously helpful in the understanding of reading disabilities. Furthermore, it is often in studying pathology that we gain our greatest insights into normal behavior. Through the study of reading/learning deficiencies, we may acquire much knowledge concerning those factors which make for interested, successful readers.

In the past fifteen years there has been a tremendous degree of interest in, and concern for, those youngsters suffering with learning disability. As might be expected, this interest has been accompanied by a good bit of controversy as well. Unfortunately, all too often the child who is experiencing learning disorder is approached with a unitary orientation so that extremely important aspects of his unique learning problem may very well be ignored. The tendency of each professional discipline to view the entire problem "through its own window of specialization" often obscures vital factors which may contribute to, or at least exacerbate, the basic difficulty. And complicating the entire picture is the fact that parents of children suffering with severe learning disabilities are ripe targets for any party who offers the "panacea." The tragic reality is that the field has become ripe pickings for anyone who wants to capitalize on the sufferings of parents who are bewildered by the multiplicity of explanations offered to them. Therefore, it becomes increasingly important for the educator to become better aware of the controversial nature of many of the diagnostic strategies and multiple intervention techniques.

THE PROBLEM OF LABELING

One of the major problems in discussing children with specific learning disabilities is the confusion of labels. Often the greatest difficulty in understanding data or claims about the different treatment approaches is that one is not sure everyone is discussing the same child. Perhaps as a reaction to the indiscriminate use of such labels as minimal cerebral dysfunction, dyslexia, strephosymbolia, perceptual handicap, minimal brain damage, etc., some educators (as well as other professional disciplines and lay people) have proposed the single unitary diagnosis of "learning disabilities." While the desire to move away from the often inaccurate "labeling" of children is praiseworthy, the conceptualization of a circumscribed area of learning disability is more than questionable. There are many different kinds of learning disorders, and each may require different types of intervention.

It would be so tempting at this point to state that we must simply abolish *all* labels that dehumanize and stigmatize both children and their families. There is no question that the whole process of labeling is fraught with many dangers. The use of noxious categorical labels with categories too narrow and too inflexible exclude many children who desire admission into certain programs. Also, labeling unnecessarily categorizes other children whose needs are diagnosed in unidemensional terms and who should not be classified by a single label. Despite all these objections, simply abolishing labels is obviously no real solution. How can we do this without taking away the very support that has allowed us to provide assistance for children with special educational needs? Historically we do know that general aid to education formulas did not benefit handicapped children until such learners were identified for purposes of legislation. This awareness has led to a proliferation of special funding for children with a wide variety of instructional needs.

DEFINITIONS, INCIDENCE, IMPLICATIONS

The Congress has defined "children with specific learning disabilities" as follows:

> those children who have a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which disorder may manifest itself in imperfect ability to listen, think, speak, read, write, spell, or do mathematical calculations. Such disorders include such conditions as perceptual handicaps, brain injury, minimal brain dysfunction, dyslexia, and developmental aphasia. Such terms do not include children who have learning problems which are primarily the result of visual, hearing, or motor handicaps, of mental retardation, of emotional disturbance, or environmental, cultural, or economic disadvantage. (Section 5(b) (4) of P.L. 94-142)

This definition of learning disabilities was originally motivated by a desire to highlight a heterogeneous group of children with developmental and academic disabilities and to provide and secure a basis for the systematic provision of financial support for training, research and service delivery. Although some of these goals have been reached, the definition has caused problems of interpretation related to operational implementation. As was stated by the Joint Committee on Learning Disabilities (1976):

1. The definition has led some to assume that learning disabilities represents a homogeneous group of children, when in fact the learning disabled con-

stitute a heterogeneous population.

2. The definition by exclusion of certain handicapping conditions as well as children from different sociocultural and linguistic backgrounds, has interfered with a clear understanding of learning disabilities as it is related to these issues.
3. Some state and local educational agencies, by a rigid interpretation of the definition, developed criteria for service delivery that ignored professional preparation and expertise and required delivery consistent with the definition as implied in the law.
4. Identification, assessment, and remediation were keyed to interpretation of the law with subsequent confusion in all these areas.
5. The lack of definitional clarity has resulted in a wide variance in agreement of the incidence of the learning disabled.

The last point related to incidence bears further elaboration. Figures on the incidence of learning disabilities within public schools range from 3 to 20 percent depending on the definition and the types of factors taken into account in deriving the figures. When learning disability is defined strictly, then the incidence is probably close to 3%. Indeed, in counting children, the Congress has stated that children with specific learning disabilities may not constitute more than one-sixth of the children counted as handicapped. Another limitation on the count is that a State may not count more than 12 percent of the number of children aged five through seventeen as handicapped. This means at most only two percent of the children in a State may be counted as having specific learning disabilities for allocation purposes.

When the prime definition of learning disability involves specific skill deficiencies which cause the child to be educationally retarded in terms of his intellectual potential, 15 to 20 percent of the school population may well be categorized as having learning problems. Within inner-city schools, when that definition has been used, incidence figures as high as 40 percent have been reported.

It is apparent that determining incidence of specific learning disability in the school population is a difficult task. All of the problems discussed above regarding definition affect the validity of attempts to count the number of learning disabled children. In addition, there is frequently confusion of etiological classifications with specific symptomatology. At times, distinctions between degrees of defect or severity of a learning disabling condition are inadequate. Finally the professional orientation of the researcher may influence the estimate, his background and viewpoint perhaps tending to influence choice of sample, etiological and symptomatic terminology, definition of the condition studied, and/or conclusion.

READING AND LEARNING DISABILITIES

One of the crucial issues in the field today is the relationship between reading and learning disabilities. Although reading difficulty constitutes only one slice of the total pie of learning disability, there can be little question that it is one of the largest slices. Reading appears to be both one of the most troublesome and most important skills in school. A knowledge of reading is necessary for the mastery of almost any school subject, while more children fail a particular grade because of difficulty in reading than for any other reason.

It should be noted that the majority of children who are classified as learning disabled have as their *major* presenting symptom difficulty in learning to read. Furthermore, at a meeting in Chicago called by the Bureau of Handicapped Children to discuss the definition and rules and regulations for P.L. 94-142, attended by experts of a variety of professional persuasions, practically every time an individual wanted to make a point about learning difficulty, he would use an example of a child with a reading difficulty. In January, 1976 the International Reading Association conducted a survey of the State Reading Requirements for Learning Disability Certification. In the majority of states, there are no reading requirements specified. In the vast majority of states, no more than one course in reading is required.

Unfortunately, a situation has arisen in this country where children who have severe reading problems cannot be taught by reading specialists because of the interpretation of the definition established by Congress. At the most recent IRA conference a resolution was adopted that "IRA conduct a vigorous campaign at the state and provincial level to acquaint legislators with the significance of the problems stated in this resolution, and to seek their support to certify graduates of reading/language specialists and as learning disability specialists." Certainly this is a position which is open to much discussion and deliberation. What seems to be most clear is that Federal and State regulatory provisions should make it clear that individuals, depending on the type and severity of their disability, may best be served by such personnel as learning disability teachers, speech and language pathologists, reading specialists, psychologists, etc. In other words, there must be real commitment to the idea that those with the competencies to remediate a particular problem should be permitted to do so regardless of the label that the teacher may possess.

DIAGNOSIS

Psychological and educational tests are familiar tools in the evaluation of children with known or suspected learning disability. Although these tests continue to play an important part in the clinical evaluation of children with learning disorder, their exact role and their utility are matters of considerable controversy. Enthusiasm for testing has its roots in the assumption that the test score gives a reliable quantitative estimate of whatever behavioral variable the test is thought to be measuring. It is easy to understand why many clinicians have come to rely so heavily upon psychologic tests in deciding whether or not a child's behavior is best explained by psychologic factors or by organic damage to the brain. After all, when other techniques are in question, one might not be blamed for turning to the psychologist to provide the ultimate answer through "testing magic." However, there is often a discrepancy between the subjective certainty of some clinical and school psychologists concerning their interpretation of test results and the frequent lack of objective information concerning the predictive power of their testing.

As is so often the case, the real value of psychological and educational testing depends to a very large degree upon the competence of the examiner. Appropriately chosen tests correctly administered and carefully interpreted constitute one more tool to study the child's behavior and attempt to understand it. In addition, special tests used in a research setting have much to offer in extending and refining presently available knowledge concerning behavior and its relation to brain function.

Despite this criticism of psychological and educational testing, the reader should not be led to believe that some of the other, more medical methods of assessment are any more reliable or valid. Even though in recent years there has been increasing evidence of the danger for the nervous system of prematurity and "reproductive casualty," Knobloch and Pasamanick (1959) reported incidents of trauma during the early years are so common that their relevance to the child's difficulty is at best conjectural. The use of "soft signs" on neurological examination is certainly open to question. It is very difficult to interpret these minor signs which are found in a significant number of presumably normal children and which some authors claim to be especially frequent in the children whom they label as brain damaged. Even electrocephalographic evidence is disputable. A grossly abnormal tracing or a reproducible focal discharge can be interpreted safely as evidence of abnormal cerebral functioning. However, the same cannot be said about lesser anomalies. Mild focal abnormalities may indicate immaturity of cerebral organization.

MEDICATION

The treatment of severe learning disabilities has involved different professional groups and will continue to be an interdisciplinary effort in the future; e.g. pediatricians, pharmacologists, psychiatrists, teachers, psychologists, and parents. Drug therapy is only one aspect of this treatment. As Eisenberg (1971) has pointed out, "Drugs promise neither the passport to a brave new world nor the gateway to the inferno. Properly employed as a single component of a total treatment program, they can be helpful in realizing the goal of the healthy development of children." The key words in that succinct and perceptive statement appear to be "properly employed."

It is important to note that stimulant drugs such as Dexadrine and Ritalin may improve the child in terms of his problems in impulse control, notably hyperactivity and hyperdistractibility. They do not affect the specific learning disabilities; hopefully, they may make the child more available for learning. It is also extremely important to differentiate between those youngsters who have neurologically-based hyperactivity and those youngsters who evidence anxiety-based increased motor activity. Finally, it is often difficult to decide whether or not a given change in behavior truly represents a specific physiologic drug effect. Indeed, possible indirect effects of the drug must be considered. The youngster's improvement may be due to his awareness of some irrelevant physiologic effect of the drug, or the drug may be completely inactive, but the child's expectation that the drug will help him may result in improvement.

There is another negative factor that must be considered. It has been asserted that reference to specific learning disability or hyperactivity as a basis for medication, is one of our most fashionable forms of consensual ignorance, and that the estimate that 200,000 children in the United States are now being given amphetamine and stimulant therapy, with probably another 100,000 receiving tranquilizers and antidepressants, represents a wholesale drugging of school-age children. There is also the expectation that the use of medication to modify the behavior of school-age children will radically increase. One must also wonder how many of these children, who are subjected to the use of such drugs as Ritalin, later turn to other drugs whenever they feel under some form of stress.

FADS AND FASHIONS

As we have suggested earlier, the field of learning disabilities has been inundated by a number of specific approaches, each of which claim to offer the cure-all for a wide variety of learning and behavioral problems. Space does not permit us to describe all of these techniques in detail, but some of the more representative will be discussed briefly.

Megavitamin therapy refers to the specific use of massive doses of vitamins to correct cases of learning disability by providing the optimum molecular environment for the mind. Cases that are usually cited as effective for megavitamin therapy are based on studies of a small group of autistic children. When one reviews the literature that has been favorable to the use of megavitamins you find references in one study to references in another study, each study citing another as the evidence for the effectiveness of the treatment. When one gets down to the final core of subjects, these are uncontrolled, non-blind studies with no objective measurements of improvement.

The best summarization of the state of the art regarding megavitamin therapy perhaps is found in a publication of the American Psychiatric Association (1973):

> ". . . in our view the results and claims of the advocates of megavitamin therapy have not been confirmed by several groups of psychiatrists and psychologists experienced in psychopharmacological research . . . The theoretical basis for megavitamin treatment especially with nicotinic acid have been examined and found wanting. The chemical and psychological tests employed for diagnosis and treatment response have been examined and found lacking in both reliability and specifity."

Treatment methods for brain damage devised by the Institutes for the Achievement of Human Potential in Philadelphia have received much publicity and have been introduced in parts of the United States and several foreign countries. The hypothesis offered by Doman and Delacato (1960) which provides the basis for the methods of the Institutes follows the principle that "ontogeny recapitulates phylogeny" and that failure to pass properly through a certain sequence of developmental stages in mobility, language, and competence in the manual, visual, auditory, and tactile areas reflects poor "neurological organization" and may indicate "brain damage." In the more severe cases of brain damage, patterns of passive movement are imposed which have as their goal the reproduction of normal activities which would have been the product of the injured brain level had it not been injured. A controversial element of the theory is that enhancement of one function will result in improvements in other areas

(e.g. gains in mobility patterns will, without special attention to speech, lead to improvement in expressive language.)

Freeman (1967) has enumerated the primary objections to the methods of the Institutes as follows: 1) a tendency to ignore the natural clinical course of some patients with brain injuries; 2) assumption that their methods treat the brain itself, while other methods are symptomatic; 3) assumption that because "the full potential of the brain" is not known, one can conclude that each child not "genetically defective" may have above average intellectual potential; 4) placing the parent in the role of therapist places on the parents the burden of possible failure of treatment, in addition to feelings of guilt and inadequacy which are probably universal with parents of handicapped children; 5) the forceful prevention of self motivated activities of the child; (6) assertions which may increase parental anxiety and concern; and 7) a variety of statistical defects.

Another great area of controversy has to do with the role of vision and visual perception in learning disabilities, specifically in relationship to dyslexia. Optometrists generally feel that learning in general and reading in particular are primarily visual perceptual tasks. They point out that visual perceptual processes are also related to sensory-motor coordination of the child. They employ a wide diversity of educational and sensory-motor perceptual training techniques in an attempt to correct educational problems in children.

While no one can argue with some of the tenets of "functional vision," and while there is no doubt that some reading problems are exacerbated by a breakdown in the accommodation-convergence relationship, one must question seriously the advisability of using visual training and visual perceptual techniques as the major method of intervention. Although the studies continue to be equivocal, a number of recent studies suggest that the relationship between training in visual perception and improvement in reading is negligible.

CONCLUSION

We have attempted to review some of the more common myths and realities of severe reading and related learning disabilities. It is clear that there have been strides in medicine, psychology, neurology, psychiatry, etc. which have contributed greatly to the knowledge of learning, thinking, and behavior. Unfortunately there still is the tendency to look at issues in biased fashion and to seek simplistic panaceas. The educator at times has been most guilty in this regard. Occasionally obsessed with the idea of specific sensory modalities as the ultimate solution for remediating the child, the teacher tends to ignore individual needs of the child which are most pertinent. Even more frequently the teacher, preoccupied with pseudoscientific approaches, forgets that he is working with a *child* — one who must be motivated and made emotionally receptive to learning.

My own experience has been that with children with learning disabilities, the most important tool in planning for alleviation of their problems is the psychoeducational programming to which they are subjected. It is true that the evaluation of psychoeducational models for these children is very difficult. The minutia of model building appears to be a favorite pasttime of some ivory-tower educators who never set foot in a classroom and who rarely talk with the parents of the children for whom they are programming. One thing is very clear: regardless of model or strategy, it is the relationship established between the child and his teacher which is of the utmost importance.

REFERENCES

American Psychological Association Task Force on Vitamin Therapy in Psychiatry. *Metavitamin and Orthomolecular Therapy in Psychiatry.* Washington, D. C.: American Psychiatric Association, 1973.

Doman, R. J., Spitz, E. B., Zucman, E., Delacato, C. H., & Doman, G. Children with severe brain injuries: neurological organization in terms of mobility. *Journal of the American Medical Association,* 1960, *174,* 257-262.

Eisenberg, L. Principles of drug therapy in child psychiatry with special reference to stimulant drugs. *American Journal of Orthopsychiatry,* 1971, *41,* 371-379.

Freeman, R. D. Controversy over 'patterning' as a treatment for brain damage in children. *Journal of the American Medical Association,* 1967, *202,* 385-388.

Knobloch, H. & Pasamanick, B. Syndrome of minimal cerebral damage in infancy. *Journal of the American Medical Association,* 1959, *170,* 1384-1387.

Recommendations of the Joint Committee on Learning Disabilities. Unpublished, April, 1976.

SHEILA K. HOLLANDER
Hollander Learning Center, Ft. Pierce, Florida

A MULTIDISCIPLINARY APPROACH TO THE ETIOLOGY AND MANAGEMENT OF LEARNING DISABILITY

The term "learning disability" is one which is bandied about in both the professional and lay literature with great frequency. Nevertheless, it is a relatively new term in the child care professions, having appeared in the literature only in the mid 1950s. Prior to that time, the inability of the child to adapt to the environment was considered primarily in psychodynamic terms. Recent medical advances, however, as well as the findings of psychological investigations have led professionals concerned with child development to a greater acceptance of the impact of the internal, physical environment on a youngster's behavior.

While the disciplines of medicine, education, psychology, optometry, and nutrition have investigated the etiology of these disabilities in learning as well as techniques of remediation, the primary care agency has been education. Many specialized techniques have been developed to cope with the symptoms of learning disability. Educators have been quite successful in alleviating a great many of these interferences with adequate learning and in assisting those afflicted with one or more aspects of this debilitating condition to cope with their problems. However, many educators have lost sight of the fact that they are dealing not with the core of the problem itself but only with its many manifestations.

For the past 10 years, many have espoused an interdisciplinary approach to this problem but few have practiced it. There has been a reluctance on the part of some groups to enter into a multidisciplinary alliance, possibly because they fear an encroachment upon their autonomy and authority. This lack of interaction has resulted in duplication of services together with the far more serious problem of omission of services needed by an at-risk population.

Many disciplines hold that the etiology of learning disability is characterized by a theory of single causation. Such a theory views a particular behavior as having its basis determined by only one set of factors. Yet familiarity with the literature of child development, normal and abnormal psychology, pediatrics, optometry and any number of other disciplines concerned with human growth and development indicates that any rationale postulating a unique treatment methodology for a particular condition is usually false.

ETIOLOGIES

Allergy

Wunderlich (1973) contended that there is a reciprocal relationship between allergy and brain dysfunction with either condition possibly being causative of the other. There is general recognition of the hereditary role played in allergy and this may bear some relationship to the genetic predisposition to learning disabilities found in many families. Hawley and Buckley (1974) claimed that hyperkinetic children exhibit an increased frequency of food allergic reactions and should be evaluated for sensitivity to food contaminants such as pesticides, fertilizer, or herbicides and for sensitivity to the various aniline coal tar dyes often found in processed foods. Feingold's work (1974) was concerned with the role of low molecular weight chemicals known as haptens. When combined with proteins, haptens may produce an allergic response. Many chemicals used as food additives and for food coloring are low molecular compounds and may function as haptens in producing sensitivity reactions. Feingold also contended that hyperactive children often display more controlled behavior when on a diet from which salicylates and artificial food colorings and flavorings have been eliminated.

Nutrition

Nutritional factors as possible causes of learning disability are of concern to many. Cravioto, DeLicardie, and Birch (1966), Crawford (1966), Penn (1966), Scobey (1947), and Thompson (1971) each stated that various nutritional deficiencies might play a role in the development of learning problems. Many believed that subsequent to early deprivation, establishment of proper nutrition is insufficient to repair the damage. Powers (1973-4, 1975) found that a treatment program limiting carbohydrates and caffeine, but including digestive enzymes and vitamin and mineral supplements was helpful in improvement of academic skills of youngsters found to have deviant blood sugar curves and symptoms of learning disability. He also noted that caffeine excites all portions of the central nervous system and that excessive amounts may cause restlessness, excitement, muscle tension, and tremulousness which are symptoms exhibited by many learning disabled youngsters. Young people who are not coffee or tea drinkers may inadvertently imbibe large quantities of caffeine through excessive drinking of sodas containing cola. Drinking of non-diet colas may result in high carbohydrate intake, which may magnify the deleterious effect of caffeine.

Brain Dysfunction

Other theories are based on malfunction of different areas of the brain. Attempts to localize lesions through electroencephalographic examination (EEG) have often proved less than reliable and its administration as part of the diagnostic procedure for the learning disabled child is considered by many to be an unproductive technique.

However, there is a growing body of research relating localized abnormalities with specific learning difficulty. Jampolsky (1970) reported that children making number reversals revealed evidence of EEG abnormality in the parietal and occipital areas. Rosenthal (1973) noted that children who respond to stimulants in what frequently has been called a "paradoxical manner" by exhibiting more controlled behavior have been found to have high electrodermal resistances and high EEG power in the 0-8 Hz frequency band which suggests a syndrome of underarousal. He further speculated (1973) that different types of dyslexias, dyscalculias, and dysgraphias are related to the hemisphere in which the lesion occurs as well as to the particular portion of the hemisphere. Nall (1973) found that EEG reports on learning disabled children indicated less than normal amounts or organization of alpha waves. In Gross and Wilson's study (1974) of 1011 child out-patients referred to a mental health center primarily for problems of underachievement, restlessness, and aggressiveness, an abnormal EEG was found in 53% although no other hard signs were noted. Frank and Levinson (1975-6) reported that in a study of over 1000 children diagnosed as having dyslexia, there was impairment of the vestibular apparatus, resulting in subclinical nystagmus leading to dyslexia. De Quiros (1976) distinguished between hyperactivity which he felt is connected with cerebral disorders and restlessness which he believed is based upon postural disorders related to vestibular, proprioceptive, and cerebellar functioning.

Biochemical Disorders

Wender (1971) hypothesized that two primary abnormalities exist in learning disabled children. These are 1) their diminished experience of pleasure and pain, and 2) an excessive and poorly modulated activity level. He theorized that the nature of these behavioral changes suggests a biochemical impairment as children thus affected have an abnormality in the metabolism of monoamines producing impairment in both the reward mechanism of the brain and the activating system of the brain. Rossi (1972) suggested that enzymatic dysfunction of genetic origin causes cerebellar malfunction and is a cause of alteration of learning processes.

Other Areas of Investigation

Mayron, Mayron, Ott, and Nations (1974, 1976), Narkewicz and Graven (1966), and Wurtman (1975) investigated the effects of incandescent and fluorescent light and radiation upon behavior. Coleman (1968), Evans, Efron, and Hodge (1976), Flax (1970), and Heath, Cook and O'Dell (1976) were concerned with the relationship of various aspects of vision to reading and learning.

MANAGEMENTS

Pharmacological

In 1937, Bradley discovered the first of the apparently paradoxical reactions produced on learning disabled children by stimulant drugs. Medications currently in use other than d-amphetamine (Dexedrine) and dl-amphetamine (Benzedrine) are methylphenidate (Ritalin) and pemoline (Cylert), the tricyclic anti-depressants, and the phenothiazines. Each group is effective with different manifestations of the syndrome and contraindicated dependent upon the side effects. Each requires close medical supervision.

Educational

The literature abounds with information on the educational management of children with learning disabilities. Hammill and Bartel (1975), Johnson and Myklebust (1967), Sapir and Nitzburg (1973) and Valett (1973) are but a few of those educators concerned with the problem whom the reader may consult for insights into the role played by education.

Psychological and Sociological

The child afflicted with learning disabilities needs more than medical and educational management. He or she needs the guidance and understanding of parents who are knowledgeable and trained to handle the cluster of behaviors often manifested. Counseling and other forms of professional assistance are frequently necessary aids for parents and siblings of the learning disabled. Gargiulo and Warniment (1976), Gordon (1975), Lupin, Braud, Braud, and Duer (1976), Luszki (1968), and Wunderlich (1975) may be consulted on the bio-psycho-social management of these children.

SUMMARY

With the exception of behavioral counseling and behavior modification the evidence implies that learning disability in its differing manifestations is of organic etiology. Nevertheless, management of this syndrome is primarily in the hands of educators who are usually lacking in knowledge of the physiological aspects of causation and the medical aspects of treatment. Similarly, medical practitioners often lack understanding of treatment procedures advocated by teachers and psychologists. There is, at best, a lack of knowledge of what colleagues in other professions are doing in this area and, at worst, a suspicion of their efforts.

CONCLUSION

Kauffman and Hallahan (1974) posit that the time has come to cease the struggle between the proponents of the medical and the psycho-educational models and to seek their unification. The contention that the medical model is concerned only with etiology while the psycho-educational model is concerned only with behavioral manifestations is a misconception of both paradigms. The traditional attitude that dealing with learning problems should be the obligation of educators alone must be replaced by recognition that handicapped children, with their various deviations, are no longer the exclusive domain of any particular discipline (Clemmens and Davis, 1969, Greene, 1969, Williams, 1976).

REFERENCES

Beery, J. W. Matching of auditory and visual stimuli by average and retarded readers. *The Optometric Weekly*, 1970, *61*, 93-96.

Bradley, C. The behavior of children receiving benzedrine. *American Journal of Psychiatry*, 1937, *94*, 577-585.

Clemmens, R. L. & Davis, J. Complimentary roles of the pediatrician and educator in school planning for handicapped children. *Journal of Learning Disabilities*, 1969, *2*, 524-532.

Coleman, H. M. Visual perception and reading dysfunction. *Journal of Learning Disabilities*, 1968, *1*, 116-123.

Cravioto, J., DeLicardie, E. R., & Birch, H. G. Nutrition, growth and neurointegrative development: An experimental and ecologic study. *Pediatrics*, 1966, *38*.

Crawford, J. E. *Children with subtle perceptual-motor difficulties*. Pittsburgh: Sanwix House, 1966.

De Quiros, J. Significance of some therapies on posture and learning. *Academic Therapy*, 1976, *11*, 261-270.

Evans, J. R., Efron, M., & Hodge, C. Incidence of lateral phoria among sld children. *Academic Therapy*, 1976, *11*, 431-433.

Flax, N. The eye and learning disabilities. *Journal of the American Optometric Association,* 1970, *41,* 841-845.

Frank, J., & Levinson, H. N. Dysmetric dyslexia and dyspraxia. *Academic Therapy,* 1975-6, *11,* 133-143.

Feingold, B. F. *Why your child is hyperactive.* New York: Random House, 1974.

Gargiulo, R. M., & Warniment, J. A parents' perspective of learning disabilities. *Academic Therapy,* 1976, *11,* 473-480.

Gordon, S. A bill of rights for parents. *Academic Therapy,* 1975, *11,* 21-22.

Greene, D. W., A study of learning problems and inter-professional relationships. In R. M. Wold (Ed.), *Visual and perceptual aspects for the achieving and underachieving child.* Seattle: Special Child Publications, 1969.

Gross, M. B., & Wilson, W. C. *Minimal brain dysfunction.* New York: Brunner/Mazel, 1974.

Hammill, D. D., & Bartel, N. R. *Teaching children with learning and behavior problems.* Boston: Allyn and Bacon, 1975.

Hawley, C., & Buckley, R. Food dyes and hyperkinetic children. *Academic Therapy,* 1974, *10,* 27-32.

Heath, E. J., Cook, P., & O'Dell, N. Eye exercises and reading efficiency. *Academic Therapy,* 1976, *11,* 425-445.

Jampolsky, G. G. Use of hypnosis and sensory motor stimulation to aid children with learning problems. *Journal of Learning Disabilities,* 1970, *3,* 570-575.

Johnson, D. J., & Myklebust, H. R. *Learning disabilities: educational principles and practices.* New York: Grune & Stratton, 1976.

Kauffman, J. W., & Hallahan, D. P. The medical model and the science of special education. *Exceptional Children,* 1974, *41,* 97-101.

Lupin, M., Braud, L. W., Braud, W., & Duer, W. F. Children, parents, and relaxation tapes. *Academic Therapy,* 1976, *11,* 105-112.

Luszki, W. A. Controlling the brain-damaged hyperactive child. *Journal of Learning Disabilities,* 1968, *1,* 44-52.

Mayron, L. W., Mayron, E. L., Ott, J. N., & Nations, R. Light, radiation, and academic achievement: Second-year data. *Academic Therapy,* 1976, *11,* 397-407.

Mayron, L. W., Ott, J. N., Nations, R., & Mayron, E. L. Light, radiation, and academic behavior: Initial studies on the effects of full-spectrum lighting and radiation shielding on behavior and academic performance of school children. *Academic Therapy,* 1974, *10,* 33-47.

Nall, A. Alpha training and the hyperkinetic child — is it effective? *Academic Therapy,* 1973, *9,* 5-19.

Narkewicz, R. M., & Graven, S. N. When children complain of fatigue. *Child and Family,* 1966, *4,* 1-6.

Penn, J. M. Reading disability: A neurological deficit. *Exceptional Children,* 1966, *33,* 243-248.

Powers, H. W. S. Jr. Dietary measures to improve behavior and achievement. *Academic Therapy,* 1973-4, *9,* 203-214.

Powers, H. W. S. Jr. Caffeine, behavior, and the LD child. *Academic Therapy,* 1975, *11,* 5-19.

Renshaw, D. *The hyperactive child.* Chicago: Nelson Hall, 1974.

Rosenthal, J. H. Neurophysiology of minimal cerebral dysfunctions. *Academic Therapy,* 1973, *8,* 291-294.

Rosenthal, J. H. Recent advances in the neurophysiology of some specific cognitive functions. *Academic Therapy,* 1973, *8,* 423-428.

Rossi, A. O. Genetics of learning disabilities. *Journal of Learning Disabilities,* 1972, *5,* 489-496.

Sapir, S. G., & Nitzburg, A. C. (Eds.). *Children with learning problems.* New York: Brunner/Mazel, 1973.

Scobey, R. R. Nutrition and child health, *N. Y.S. Journal of Medicine,* 1947, *47,* 1786-1789.

Thompson, H. L. Malnutrition as a possible contributing factor to learning disabilities. *Journal of Learning Disabilities,* 1971, *4,* 312-315.

Valett, R. E. *Learning disabilities: diagnostic-prescriptive instruments.* Belmont, Ca.: Lear Siegler, 1973.

Wender, P. H. *Minimal brain dysfunction in children.* New York: Wiley-Interscience, 1971.

Williams, J. F. Hang together or hang alone? *Academic Therapy,* 1976, *11,* 393-395.

Wunderlich, R. C. *Allergy, brains, and children coping.* St. Petersburg, Fla.: Johnny Reads, 1973.

Wunderlich, R. C. Biosocial factors in the child with school problems. *Academic Therapy,* 1975, *10,* 389-399.

Wurtman, R. J. The effects of light on man and other mammals. *Annual Review of Physiology,* 1975, *37,* 467-483.

CATHERINE MORSINK
University of Kentucky

WAYNE OTTO
University of Wisconsin-Madison

SPECIAL CONSIDERATIONS IN TEACHING LD CHILDREN TO READ

Reading disability is the main problem of many children labeled, "learning disabled." Samuels (1970) defines reading disability in this way:

> "Generally, a reading disability is said to exist when despite adequate instruction, absence of emotional problems which may interfere with learning, a cooperative child, and absence of sensory impairment, there is a discrepancy between the child's reading achievement level and some measure of potential ability." (p. 267)

The part of Samuels' definition with which the present paper is most concerned is the phrase . . . "despite adequate instruction . . ." One of the present authors (Otto & Chester, 1975) has suggested that the assumption of adequate instruction, in many cases, may be false, and that educators ought to be primarily concerned with educational factors as causes of reading failure.

The study described in this paper was based on the assumption that more adequate instruction can increase the reading levels of disabled learners, and that one component of adequate instruction is an organizational framework which helps learners to focus on what they need to know. *The Wisconsin Design for Reading Skill Development: Word Attack* (Otto & Askov, 1972) provides such a framework. It includes identification of essential content (skills), statement of objectives, assessment, identification of appropriate teaching/learning activities, and evaluation.

The Design was developed for use with normal learners. Yet, use of this skill-centered framework with disabled learners seems to offer some advantages: its pretesting mechanism saves teaching time; its focused instruction centers directly on the child's skill deficits; its mastery emphasis provides as many opportunities for learning as necessary; its tracking system facilitates communication with other professionals.

Some evidence on the effectiveness of the assessment-evaluation components of the framework with slow learners has begun to accumulate (Rosencranz, 1973; Mullen, 1976; Morsink & Otto, in press.) To date, there have been no reports that the teaching/learning activities (the Teacher's Resource File of *Design: Word Attack*) have been used experimentally with disabled learners. Since instruction designed for normal learners may or may not be appropriate for disabled readers, the total *Design: Word Attack* framework, which includes the teaching/learning activities, was subjected to testing in the present study.

The major research question for the study was as follows:

1. Will reading achievement (decoding words in isolation) differ among groups of disabled learners assigned to instruction organized around the *Design: Word Attack* framework and two control treatments?

In addition, because the presence of severe learning difficulties indicated that measures of effectiveness should include assessment of the degree to which disabled learners can retain and apply skills, two additional research questions were asked:

2. Will ability to apply skills to oral reading in context differ among the groups?
3. Will experimental students retain learned skills after the passage of time and the intervention of other instruction?

The final objective of the study was to determine whether changes in the suggested teaching/learning activities might increase the utility of this instructional framework for exceptional children.

METHOD

Sample

Subjects were seven-through-twelve-year old disabled learners (Kentucky state guidelines) with extreme difficulty in word attack. There were four to nine subjects in each of ten LD classrooms and all ten teachers taught students in all three treatments. An analysis of variance indicated that groups assigned to the three treatments were not equivalent in terms of pretest achievement. An analysis of covariance, using pretest scores as the covariate, was therefore selected as the appropriate technique for analysis.

Data Collection and Instrumentation

Reading words in isolation. The Wide Range Achievement Test: Reading (WRAT:R, Jastak and Jastak, 1965) was used to assess students' ability to decode words in isolation. The WRAT:R was administered as a pretest in September and as a posttest in January.

Application: Reading words in context. Students' ability to read words in context was measured by charting the number of correct words per minute read orally in a basal reader at instructional level.

Retention of skills. Retention of skills was measured, for the experimental group only, by re-administration in May of the alternate form of the *Design: Word Attack* tests over skills each student had mastered during the year.

Documentation of needed changes. Structured interviews with all ten teachers were used to document recommended changes in the teaching/learning activities currently suggested in the Design's Resource File. In addition, classroom observations of the performance of the ten teachers and their thirty-one student teachers were used to document the presence or absence of selected teacher behaviors which might have an impact on students' learning. Both the interview and the observational techniques are described in more detail elsewhere (Morsink, 1976a, 1976b).

Experimental and Control Treatments

Students were assigned to three treatments: basal reader plus *Design: Word Attack* (Experimental), basal reader only (Control - 1), and programmed instruction featuring spelling patterns (Control - 2). All three treatments featured overall instruction in a basal reader. For the Experimental and the Control - 1 groups, the basal reader for most students was the Houghton-Mifflin series (Durr, La Pere, & Niehaus, 1974). For the Control - 2 group, the major basal reader was *Programmed Reading* (Sullivan and Buchanan, 1963). The unique attribute of the Experimental treatment was its provision of focused word attack instruction, using the Design's diagnostic-prescriptive framework.

RESULTS

The analysis of covariance on the January WRAT:R posttest scores, using the September pretest as the covariate, indicated that the effects of treatments were significantly different, $F(2, 49) = 4.52, p = < .02$. Adjusting the January means by the pretests, the Experimental group had a grade level mean of 3.05, followed by the Control - 1 group mean of 2.69 and the Control - 2 group mean of 2.47.

Graphs of students' ability to read words in context indicated that context reading for the Experimental group (pretest = 47.2 WPM, posttest = 55.9 WPM) was superior to that of the other two treatments (Control - 1, pretest = 49.1 WPM, posttest = 49.6 WPM; Control - 2, pretest = 34.4 WPM, posttest = 36.6 WPM).

The results of Design posttests, re-administered in May for the sixteen students remaining in the Experimental group for the entire year, indicate that in fifteen of the sixteen cases, students retained all or all but one of the skills learned. There were, however, large differences in the number of skills mastered by various students. Some of the variation in students' performance was apparently related to the classroom in which they were taught. In one of the classrooms, students mastered an average of only one word attack skill, while in another classroom students averaged nearly seven new skills.

Those teachers whose students made greatest progress reported spending a great deal of time designing their own teaching/learning activities. Results of teacher interviews indicated that the existing Teacher's Resource File did not provide an adequate number of teaching ideas and that references were of little assistance because the suggested materials were either unavailable or inappropriate for exceptional learners.

DISCUSSION

The results of this study should be interpreted with caution, since two of the data collection techniques were experimental in nature. Neither the procedure for measuring oral context reading, nor the technique for observing teacher behavior has been validated. In addition, the number of subjects available for randomized assignment to treatments was admittedly small. The reader should also note that this experiment focused on word attack skills and did not include a consideration of reading comprehension.

In this study, students in the Experimental group demonstrated superior achievement for reading words in isolation and they were able to retain learned skills. Students in the Experimental group also had higher rates of accurate oral context reading, as measured by an experimental technique. These results lend some support to the premise that the *Design: Word Attack* may provide an appropriate instructional framework for teaching disabled learners to read. The aspects of this framework which may be of particular value seem to include its emphases on focused instruction and on mastery learning. This inference is consistent with the findings of two earlier studies in which skill-centered instruction was evaluated for disabled readers (Guthrie, 1974; Morsink, 1974).

Not all LD children in the study responded equally well to the experimental treatment, and not all teachers implemented experimental procedures equally well. Data gathered during classroom observations and teacher interviews suggest that the way in which teachers used the materials was an important consideration in whether students learned successfully.

Teacher feedback, classroom observations, and data collected during this study indicate that, for disabled learners, the strength of the Design's framework seems to be its diagnostic/prescriptive system for providing focused instruction, while the weakness is its suggested teaching/learning activities (Teacher's Resource File).

Because of this finding, a new Design program, entitled DELTA: Word Attack Adaptation for Disabled Readers, was developed. The new instructional materials provide detailed teaching sequences for presenting instruction that is focused on skill development. The materials also provide for instruction which is adapted to meet the needs of disabled learners by incorporating special features and including suggestions for teacher use. The special features of these new materials are based on an analysis of disabled readers' special learning needs and they are described in more detail in the rationale for the new program (Morsink & Otto, in press).

Confusion over the definition and treatment of reading/learning disability will probably continue for some time. Many researchers suggest that there are different kinds or degrees of reading disability, and that there is a continuum of disability. If this is true, then classification of students should follow, rather than precede, treatment and be based on the child's response to treatment. No child should be classified as a disabled learner until there is evidence of adequate instruction.

REFERENCES

Durr, W.; La Pere, J.; & Niehaus, B. *Houghton-Mifflin readers* (2nd ed.). Boston: Houghton-Mifflin Company, 1974.

Guthrie, J. *Identification and instruction of children with reading disability.* Second Annual Report to the Spencer Foundation, July, 1974. (ERIC Document 098516).

Jastak, J., & Jastak, S. *Wide Range Achievement Test.* Wilmington, Delaware: Guidance Associates, 1965.

Morsink, C. *Reading: An educational approach to disability.* Lexington, Kentucky: University of Kentucky RSEIMC, 1972.

Morsink, C. *A study of the effectiveness of the READ materials on the growth in reading achievement of second graders in a compensatory reading program.* Unpublished doctoral dissertation, University of Kentucky, 1974.

Morsink, C. *Some teacher competencies that may influence disabled readers' learning of word attack skills.* Lexington, Kentucky: University of Kentucky, Department of Special Education, 1976a.

Morsink, C. *Utility of the Wisconsin Design Word Attack system for exceptional children: A summary of teacher's perceptions.* Lexington, Kentucky: University of Kentucky, Department of Special Education, 1976b.

Morsink, C., & Otto, W. *DELTA: Word attack adaptation for disabled readers. Teacher's planning guide, part one: Rationale.* Minneapolis, Minnesota: National Computer Systems (in press).

Mullen, D. *Summary of project COMP, 1975-1976.* Marinette, Wisconsin: Marinette Public Schools, 1976.

Otto, W., & Askov, E. *The Wisconsin Design for Reading Skill Development: Word Attack.* Minneapolis, Minnesota: National Computer Systems, 1972.

Otto, W., & Chester R. *Objective-based reading:* Reading, Massachusetts: Addison, Wesley, 1975.

Rosencranz, C. EMR reading program also helps regular students. *Wisconsin Journal of Public Instruction,* Autumn, 1973, 23-28.

Samuels, S. J. Reading disability? *The Reading Teacher,* 1970, *24,* 267.

Sullivan, M. W., & Buchanan, C. *Programmed Reading.* New York: McGraw-Hill, 1963.

PATRICIA M. CUNNINGHAM
Alamance County Schools, N.C.

INVESTIGATING THE ROLE OF MEANING IN MEDIATED WORD IDENTIFICATION

Historically, word identification has been viewed as a process which, when successfully completed, resulted in the gaining of some meaning for the reader. Recently, psycholinguists and others in the field of reading (Smith, 1971, for example) have asserted that meaning is not only a product of the word identification process but also an essential component of that process. The reader, when confronted with an unfamiliar word, applies knowledge of letter-sound relationships and then compares the resulting word to words for which the reader has an established acoustic category. When first confronted with the word *crow,* for example, a reader might arrive at a pronunciation which rhymed with the word *how* or one which rhymed with the word *grow.* Pronouncing *crow* as if it rhymed with *how,* however, does not produce a word which the reader has heard before (acoustic match). For most American Readers, pronouncing *crow* so that it rhymes with *grow* does result in a word for which the reader has an acoustic and a semantic match. Often the reader must use the context of what he or she is reading to determine the correctness of a pronunciation. The word *lead* for example may rhyme with *seed* or with *bread.* Only the semantic absurdity of pronouncing *lead* to rhyme with *bread* in the sentence: *Right up to the finish line, David was in the lead,* encourages the reader to try to arrive at an alternative pronunciation. Achieving an acoustic and semantic match gives feedback to the reader that he or she has successfully completed the word identification process.

The use of nonsense words in reading tests which purport to measure a student's word identification ability is almost a standard procedure in most reading diagnoses and a common component of most reading tests. (Guszak, 1972; Gates and McKillop, 1962; and Spache, 1963. Furthermore, nonsense word pronunciation tasks are commonly used in empirical research in which the word identification process is investigated (Rosinski and Wheeler, 1972; Hardy, Stennett and Smythe, 1973; and Railsback, 1970). The results of research by Cunningham (1975-76) suggest that performance on a nonsense word pronunciation task may not be indicative of subjects' word identification ability. In that study, second graders' word identification skills were assessed using both nonsense words and low-frequency real words. No experimental-control group differences were apparent on the nonsense word pronunciation task, but significant differences were apparent on the low-frequency word pronunciation task. Had only the nonsense word pronunciation task been used as a measure of the success of the word identification treatment, the treatment would have been wrongly assessed as ineffective.

The purpose of the present study was to further investigate the role of meaning as a mediator in word identification and to assess the efficacy of nonsense word pronunciation tasks in diagnosis and research to assess word identification abilities.

METHOD

The subjects of this investigation were 69 second graders enrolled in a southeastern suburban public school. Data was collected in November when the subjects had been in school approximately three months. In order to determine that the two-syllable words used in this investigation were not sight words for this population of second graders, a random 20 percent sample of the 69 second graders was selected. The investigator tested each of these 13 children individually by displaying each of 15 two-syllable words for one second. The results of this flash test confirmed the fact that the 15 two-syllable words were not sight words for the second graders. (Eleven children correctly recognized none of the 15 words, one child correctly recognized two words and one child recognized 12 words. The mean number of correctly pronounced words was less than one. The median and mode were both zero.)

The remaining 80 percent of the population were then randomly assigned to two groups. One group was designated as the "real-word group" and the other group as the "nonsense word group." Members of the real-word group were shown cards on which were typed the 15 two-syllable words shown not to be sight words for most second graders. These 15 words were typed so that the first syllable of the word was separated from the second syllable by five letter spaces. Children in this real word group were given four examples and then told to try to figure out the 15 real words. Subjects were further told that "the two parts put together will make a word you know. Try to figure out each part and come up with a word you know." (The 15 two-syllable words used were: *fever, umpire, stampede, survive, bonus, debate, termite, clover, harness, problem, darling, vibrate, temper, suspect* and *cactus.)*

Members of the nonsense-word group were shown cards on which were written nonsense words. (The nonsense words were constructed by randomly matching a first syllable from one of the 15 two-syllable words with a second syllable. The 15 nonsense words were: *febate, umper, stamlem, surnus, bover, depede, tertus, clobrate, harling, probmite, darver, viness, tempect, suspire,* and *cacvive.*) These 15 nonsense words were also typed so that the first syllable was separated from the second syllable by five letter spaces. Members of the nonsense word group were given four examples and were told: "The two parts put together will not make a word you have ever heard of. I made up these words so no one has ever heard them before. Try to figure out each part and come up with a nonsense word — a word you have never heard before."

Reliabilities for the real-word and nonsense-word pronunciation tasks were computed using the KR_{21} formula. For the real-word task, the reliability was .93; for the nonsense-word task, reliability was .91. Subjects' responses to both tasks were scored according to the number of syllables correctly pronounced.

One of the assumptions of this experimental design was that the 15 real words were indeed words which these second graders had heard before. While it was not essential that these second graders had a meaning stored for these words (semantic match), it was essential that they recognized the word as a word they had heard (acoustic match). In order to determine if the 15 real words were words for which these second graders had an acoustic match, a second task was administered to all subjects one month after the administration of the word pronunciation task.

For the second task, all subjects were presented with a randomly-ordered list of the 30 stimuli (real words and nonsense words) used in the word pronunciation task. Subjects listened as the investigator read the stimuli and then circled those which were real words and crossed out those which were nonsense words. Before beginning this task, the subjects completed eight sample items together. Their responses to these items indicated that they understood the nature of the task they were being asked to complete. Their responses to the task also indicated their comprehension of what they were asked to do. Of the 15 nonsense words, the subjects correctly identified an average of 14.2 as nonsense words. Of the 15 real words, 12 were correctly identified as real words by at least 80 percent of the subjects. The other three were correctly identified by less than 35 percent of the subjects. The syllables of these three words (*debate, stampede* and *harness*) were not included in the scoring of either the real or nonsense word task.

RESULTS

Subjects in the real-word group correctly pronounced an average of 13.42 syllables ($SD = 7.29$, $N = 26$). Subjects in the nonsense-word group correctly pronounced an average of 10.12 syllables ($SD = 6.73$, $N = 26$). Results of a t test, $t(50) = 1.70$, indicated that the difference between the means was significant beyond the .05 level. The number of syllables correctly pronounced by the real-word group was significantly greater than the number of syllables correctly pronounced by the nonsense-word group.

DISCUSSION

The results of this experiment lend support to the theoretical position that meaning is not only a product of the word identification process but is also a crucial component of that process. The reader, in decoding an unknown word, appears to have a mind set to produce something that sounds like a word he has heard before. The achievement of an acoustic match gives feedback to the reader that he has successfully completed the decoding process.

The results of this experiment along with the results of an earlier experiment (Cunningham, 1975-76) raise serious questions about the use of nonsense word pronunciation tasks to assess word identification abilities in both diagnostic and research settings. If meaning is a mediator in word identification, results based on a purposefully meaningless task cannot be indicative of the reader's word identification ability. Diagnosticians and researchers who have a need to assess the word identification abilities of their subjects must use a task which allows meaning to function as a feedback system for the word identification process. A more valid assessment of the subjects' word identification skills would be to flash real words and then, for analysis purposes, present the words not identified. (Durrell, 1955, uses this procedure to assess word analysis skills.) Test constructors should take this into account when revising their word identification tests and reading educators and diagnosticians should consider the role of meaning as a mediator when determining which tests to use and to recommend to teachers.

REFERENCES

Cunningham, P. M. Investigating a synthesized theory of mediated word identification. *Reading Research Quarterly,* 1975-76, *11,* 127-143.

Durrell, D. D. *Durrell analysis of reading difficulty.* New York: World Book Company, 1955.

Gates, A. I. & McKillop, A. S. *Gates-McKillop reading diagnostic tests.* New York: Teachers College, Columbia University, 1962.

Guszak, F. J. *Diagnostic reading instruction in the elementary school.* New York: Harper and Row, 1972.

Hardy, M., Stennett, R. G. & Smythe, P. C. Word attack: How do they figure them out? *Elementary English,* 1973, *50,* 99-102.

Railsback, C. E. Consonant substitution in word attack. *The Reading Teacher,* 1970, *23,* 432-435.

Rosinski, R. R. & Wheeler, K. E. Children's use of orthographic structure in word discrimination. *Psychonomic Science,* 1972, *26,* 97-98.

Smith, F. *Understanding reading.* New York: Holt, Rinehart and Winston, Inc., 1971.

Spache, G. D. *Diagnostic reading scales.* Monterey, California: McGraw Hill, 1963.

MARTHA C. EVANS
University of Maryland

NANCY E. TAYLOR
Catholic University of America

IRENE H. BLUM
University of Maryland

FUNCTIONAL WORD IDENTIFICATION ABILITY OF POOR READERS

Children who perform poorly on school related reading tasks are often assumed to have no reading skills. It is not uncommon for teachers and diagnosticians to refer to these children as "non-readers". It may be that this perception of reader ability reflects a narrow view of what reading is. As society becomes increasingly visually complex, children encounter print in a wide variety of meaningful, functional situations. They are exposed to words in a variety of contexts ranging from street signs to sophisticated commercial advertisements. Young children who might be considered "pre-readers" by traditional standards frequently recognize the names of products they like. Many students who are considered failures in school demonstrate good ability to function out in the world in spite of their inability to decode and interpret print encountered in books.

Elementary school reading programs typically use basal readers as the primary source of instructional materials. Children are taught how to decode and interpret print by instruction in a series of skills which, hopefully, give them strategies for analyzing, identifying and responding to written language. Success in the school reading program is usually defined in terms of steady progress through the reading series and high standardized tests scores. Children who are learning to read are not, however, totally naive in terms of strategies for extracting meaning from symbolic aspects of their environment. Since birth they have been exposed to advertisements, labels, street signs, billboards, etc. along with books and magazines. They have dealt with these written materials in many "real world" situations. The strategies children use in these real world reading encounters are deduced after meaningful exposure to visual materials. They are not developed by instruction in specific skills. Little attention has been paid to the functional reading skills of young children nor to the real world reading ability of children who have not been successful in the school reading program.

One of the crucial differences between reading material encountered in school texts and tests and reading material encountered in the real world is that there is more non-linguistic information available in the functional situation. Entwistle (1971) has suggested that some children develop "styles of learning" at home which may differ from those they meet in school, and cites the processing of non-linguistic information in the environment as a variable on which beginning readers may differ. It is also possible that the use of visual non-linguistic information in the environment is a variable on which beginning readers may differ. It is also possible that the use of visual non-linguistic information may differ because of the other cues available in a situation.

Mason (1965) attempted to determine whether elementary aged children learned to identify words which they saw frequently on television. Words which children had been exposed to during their prime viewing time were written to resemble the televised presentation and he found that correct responses ranged from 10% for the first graders to 85% for the seventh graders, with second graders identifying 48% correctly. In addition, the study also revealed differences in the number of words superior, average and poor readers learned in this incidental fashion.

Television is, however, only one of the many ways in which children encounter print in their daily lives. Current theory holds that words are identified on the basis of their distinctive features; those aspects of the stimulus array that serve to differentiate one item from another (Gibson, 1967). The distinctive feature one needs to detect in identifying words may be a function of the situation in which these words are encountered (Smith, 1971). Children may pay little attention to print encountered in television commericals because the audio link is available, but may attend more to the graphic information on labels and signs, for example. Or the visual context cues available in all functional reading situations may eliminate attention to the distinctive graphemic features which would allow transfer of that real world word identification to the same word presented in typical print.

This study attempted to discover whether children who are considered to be either average or poor readers have a corpus of words encountered in their everyday environment that they can read, and to identify the influence of non-linguistic information available on correct word identification. This non-linguistic information was considered to be print style and the inclusion of other supporting information such as color, pictures, environment, etc. that accompanies the target word when it is encountered in daily life.

METHOD

Instrument Construction

A test was needed to assess children's functional reading skills, their ability to read words encountered in their everyday experiences. To generate a pool of valid items a group of children, from the same geographic areas as those used in the study, were interviewed to identify: (1) favorite foods, (2) familiar traffic signs and store names, and (3) television shows and commercials which they watched. The names of items most frequently mentioned in each category were incorporated into a Functional Word Identification Test (FWIT) with fifteen items in each of the three categories.

Functional word identification ability was operationally defined as the ability to correctly label the test stimuli in any of three context or cueing conditions. These conditions were:

1. typical print (items were presented in primary type)
2. stylized print (the distinctive lettering used by the manufacturer or advertiser)

3. total cueing (the actual label for food, background pictures for signs and the television items)

Food and sign items appeared on individual 5 x 8" tag cards. The television shows and commercials were seen on video tape which included the introductory video portion in the total cueing condition but not any sound. For example, Sears and McDonalds were first seen typed in primary type, second in the distinctive lettering used in the advertising. In the third condition, a photograph of Sears and the end of the McDonald's television commercial were presented.

Subjects

The subjects for the study were 48 second grade students, none of whom had repeated any grades. No attempt was made to match them on any other characteristics. They were randomly selected from four available populations, three groups of poor readers and one of average ability. Poor readers were available groups of children viewed as poor readers by their school systems on the basis of their low performance on standardized tests and school reading tasks. Two of the three groups were receiving instruction in special classes; one in an urban clinical school for disabled readers, and the other in a special class for children with specific language/reading difficulties in a suburban school. The third group received instruction in a regular class setting; these children were from the lowest tracked class having been homogeneously grouped on the basis of their low standardized test scores. Average readers were children who were receiving instruction in regular classrooms in the suburban school. These average readers were selected from the classrooms the suburban poor readers would have attended if they had not been identified for special placement.

Procedures

The Functional Word Identification Test was administered individually to each child by three examiners. A training session was conducted to insure that all items were presented in a standard manner. Items were presented one at a time in the three cueing conditions. A response was elicited in each cueing condition. The examiners did not respond to the correctness of previous responses but for each stimulus asked the child "What do you think this says?" Total administration time was twenty to twenty-five minutes per child. Criteria for judging correct responses were established on the basis of how an able reader would identify the stimuli. Responses which simply identified the picture or the category of the item (i.e. candy, a restaurant) were considered incorrect.

Data Analysis

Scores on the Functional Word Identification Test were analyzed by analysis of variance procedure. A 4x3x3 repeated measures design was used. The factors were reader (poor readers in a clinical setting, in a special class, in homogeneously grouped low track class, and average readers); amount of cueing (typical print, stylized print, and total cueing); and type of item (food labels, signs and store names, and TV shows and commercials).

RESULTS

The means and percentage of items of the Functional Word Identification Test correctly identified by each group of readers in each cueing condition are presented in Table 1. When the only cue was the word in primary type the percentage of words identified by the three groups of poor readers ranged from 19 to 24 percent. When the distinctive printing style was available as a cue the percentages of correct responses ranged from 30 to 39 percent. When all non-linguistic information was available these percentages ranged from 59 to 68 percent correct. The average readers' performance exceeded all the groups of poor readers under all cueing conditions.

The results of the repeated measures analysis of variance showed that the main effect of cueing was significant, $F\ (2, 44) = 318.99, p < .0001$, and there were significant interactions between cueing condition and type of reader, $F\ (6, 44) = 3.851, p < .01$, and between cueing condition and item type, $F\ (4, 44) = 97.39, p < .0001$. The results of the Newman Keuls tests of the Reading x Cueing interaction indicated that the performance of poor readers on item presentation in primary type was significantly lower than the group of average readers, although there were no differences among the groups of poor readers. When the words were presented in stylized print, thus preserving their distinctive graphic characteristics, the average reader's performance was not significantly different from the low-track class but was significantly higher than the clinical and special classes. Table 1 includes the relevant means.

Table 1

Mean Scores and Percentages for Functional Word Identification Test for Cueing Conditions and Types of Reader

	Cueing		
Type of Reader	Typical Print	Stylized Print	Total Cueing
Poor Reader Clinical Setting	9.00 20%	13.58 30%	26.75 59%
Poor Reader Special Class	8.92 19%	14.33 32%	29.42 65%
Poor Reader Regular Class	10.67 24%	17.50 39%	30.83 68%
Average Reader Regular Class	20.30 45%	24.08 54%	32.75 73%

A Newman Keuls test of the significant item by cueing interaction indicated that performance across all groups was higher on sign items than on label and television items in both the print and stylized print conditions. However the performance pattern was reversed in total context where the sign item mean was the lowest of the three item types. Table 2 presents the relevant means.

Table 2

Mean Scores and Percentages for
Item Type Under Three Cueing Conditions

	Cueing		
Item Type	Typical Print	Stylized Print	Total Cueing
Labels	3.00	5.23	10.56
	20%	35%	70%
Signs	6.02	7.21	8.02
	40%	48%	54%
TV	3.08	4.94	11.35
	20%	33%	76%

DISCUSSION

The results of this study indicate that children who are disabled readers in school situations can do functional reading when total context is available. Although these children had all demonstrated sufficient difficulty on school reading assessment measures, when they encountered functional words in a meaningful context their performance was similar to that of average readers. The distinctive feature theory (Gibson, 1967; Smith, 1971) holds that one uses only those cues necessary to distinguish one object from another. When the reading context for functional words resembles the context in which these words are encountered in real life there are no differences between children classified as disabled readers and those classified as normal readers. All groups seemed to be equally able to attend to those cues necessary for correctly identifying many of the test stimuli used in this study, although these features may not all be of a graphic nature.

Different response patterns were noted however, among the disabled readers' approach to the total context items. For some children total cueing meant interpreting the picture for they supplied a generic label that did not indicate attention to the graphically provided particulars of the item. An example is identifying *Swanson's Fried Chicken* as frozen dinner or TV dinner. For others, the focus was clearly on the print to the extent that some aspects of the non-linguistic context were ignored. This was particularly apparent in the TV segments where for example, one child saw the space ship picture lead-in behind the words Star Trek but read the title as "Stan Turkey."

The ability to utilize the cueing in print and stylized print conditions did differentiate between the suburban average readers and the two groups of poor readers in special classes. Suburban average readers were able to read significantly more items when they occurred in typical print and in the highly specific stylized print used by manufacturers and advertisers. There are two possible explanations for this finding. It may be that the graphic cues are not the salient cues when children encounter these items in their everyday experiences. There may be no need to attend to graphic aspects because cues of color, arrangement, label, shape, pictures, and the environment serve to differentiate items. A related argument is that success with, or instruction in, the graphic code serves to draw attention to the graphic aspects of "real world" linguistic displays. Thus suburban average readers may have been able to perform significantly better than some groups of poor readers in the print and stylized print conditions because they were successful with the code.

All groups performed differently on items based on signs than on the food label and television items. Performance on these sign items was higher in the print and stylized print conditions but there was less improvement when the stylized print was supplemented with total context. Words on the signs (stop, one way, on, giant, etc.) are words children are likely to encounter in school as well as in "real world" situations. Therefore it is possible that this higher degree of exposure accounted for higher performance on these items in the print and stylized print conditions. There is also less additional information available as one moves from stylized print to total cueing (e.g. no color was added and the print of these signs in the stylized condition does not differ from the print used in the print condition or when the total sign is exposed).

IMPLICATIONS

Downing has stressed that children approach reading instruction in a state of cognitive confusion (1972). They do not understand what one does when one reads or the purpose the act of reading serves. Using the content of the child's environmental encounters and recognizing the strategies he has developed naturally from these encounters may serve to foster cognitive clarity. Such application also serves to reinforce the purposeful nature of encoding print, a message that is sometimes lost in certain types of classroom instruction.

The results of this study suggest that our traditional measures of reading ability may be ignoring one aspect of children's reading performance: their ability to read in a real world situation. Although standardized tests may correctly identify children who are not succeeding in a school's reading program and who may profit from instructional modification, the fact that groups of similar children perform very differently in reading in school does not imply that their functional "real world" reading abilities differ drastically. While the classifications of poor and average readers may be appropriate for instructional activities, these children are equally able to buy the right candy bar at a store whose name they recognize.

The challenge lies in finding ways to facilitate the transfer from what the child already knows to the more generalizable aspects of graphic coding (i.e. the orthographic system). The distinctive characteristics of environmental items are highly salient but they are also highly item specific. The salient features of color and script in Coca Cola may not generalize to other words. They may, however, be a beginning place for helping children develop successful strategies for dealing with all printed material.

REFERENCES

Downing, J. The cognitive clarity theory of learning to read. In V. Southgate (Ed.), *Literacy at all levels.* London: Ward Lock, 1972.

Entwistle, D. R. Implications of language socialization for reading models and for learning to read. *Reading Research Quarterly,* 1971, *6,* 111-167.

Gibson, E. *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts, 1967.

Mason, G. E. Children learn words from commercial television. *Elementary School Journal,* 1965, *65,* 316-320.

Smith, F. *Understanding reading.* New York: Holt, Rinehart & Winston, 1971.

JOSEPHINE S. GOLDSMITH
Rutgers University

MARK J. NICOLICH
Rider College

WORD BOUNDARIES REVISITED: A FIRST GRADE STUDY

The word boundary task (Klein & Klein, 1972) has been used to explore the effects of context on individual word identification decisions. Workers with the task have attempted to use it to infer the preferred syntactic unit size utilized in stages of developmental reading (Klein, Klein and Bertino, 1974; Goldsmith, 1974). It has been posited that the point at which larger units (phrases, sentences, full discourse) facilitate individual word boundary decisions may signal a stage at which these units are optimal for the readers being studied.

The word boundary task consists of strings of words printed in lower case with a single space between each letter, no additional spaces between words and all capitals and punctuation omitted. The subject's task is to slash boundaries between words so as to leave no extra letters. Time is controlled so that the subject who makes incorrect decisions, leaving extra letters unaccounted for, is penalized by having to go back to correct his decisions.

Following are two samples of word boundary passages (Goldsmith, 1975), one in random word format, the other presenting full discourse.

Sample 1 *(Grade 4, Hard Words)*

w a s h a v e f o r j o h n n y
t o o h e t o m o t h e r e m p t y
s l e e p e a t i t s a w f u l
i n s i d e f e e l t o i m g e t

Sample 2 *(Grade 4, Hard Discourse)*

l i t t l e j o h n n y m o s e s
s a t u p i n h i s t r u n d l e
b e d i m a w f u l h u n g r y
a n n i e h e s a i d s o f t l y

A reliable finding in word boundary work has been that context is a powerful facilitator of word boundary decisions in adults (Klein & Klein, 1972; 1973; Klein, Klein & Hildum, 1974; Klein, Klein & Vigoda, 1972) and with children in grade 4 and above (Klein, Klein & Bertino, 1974; Klein, Klein & Doris, 1973).

However, attempts to demonstrate effects of particular syntactic units such as phrases or sentences have not been successful. With the exception of one study (Klein, Klein & Hildum, 1974) which showed usefulness for articulatory phrases, results tend to support a shift from word by word reading to utilization of full discourse.

The location of this shift has itself been elusive. The Klein studies find little effect of context at grades 3 and below. However, a cross sectional study of performance in grades 2, 4, and 6 (Goldsmith, 1974) found full context facilitated correct decisions in upper SES children as early as grade 2.

The present study investigated the performance of first grade children on two versions of the Word Boundary Task: random words versus full discourse.

METHOD

Subjects

Sixteen children of each sex were selected at random from two intact first grade classes of an elementary school in central New Jersey. Within each sex group half of the children were assigned to the Random, half to the Discourse treatment condition.

Material

Passages had been selected from elementary reading materials at the Reading Center, Rutgers Graduate School of Education. Four passages were chosen, two at each of two difficulty levels: Easy (first grade level) and Hard (second grade). Passages were calibrated through the agreement of two readability formulas: Fry and Spache. Each passage of 60 words was prepared in two versions: Random words and Discourse. Discourse passages began at the first word of a selection or story. Since these passages had been used in a previous study (Goldsmith, 1974), where a statistical test for a passage effect proved non significant, they were accepted as providing suitable alternate forms.

Procedure

Children were tested within the class to provide minimal disruption. All children were given booklets and assumed themselves to be part of the study. Protocols of children whose names did not come up in the selection were discarded. Equal numbers of males and females were selected for each chunk (random versus discourse) level. Each child received four 60 word word boundary tasks: two replications for each difficulty level, easy and hard.

Design and Analysis

A 2 between/2 within repeated measurements design was utilized to assign subjects to treatments and for the analysis of variance. The between factors were sex and chunk (random or discourse). Repeated measurements were taken for difficulty (hard or easy) and for passages (a replication) nested within difficulty.

RESULTS

The analysis of variance revealed no significant main effects and only one significant interaction, chunk by passage (within difficulty), $F(2, 56) = 4.51, p < .01$.

Table 1 includes the means involved in the chunk by passages (difficulty) interaction. As can be seen from the differences column, the differences between passages are small and insignificant for the random word condition, while they are large and significant ($p < .01$) for the discourse condition. Furthermore, the direction of differences is reversed (compared to the other three) for the two passages in the hard discourse condition. Apparently, there were some syntactic factors operating within easy passage 1 and hard passage 2 that particularly interfered with children's ability to make word boundary decisions. An informal inspection of these two passages yields no obvious support for such an interpretation, however.

Table 1

Means for Word Boundary Scores by Chunk
Difficulty and Passage

Chunk	Difficulty	Passage	Means	Difference
Words	Easy	1	15.56	2.63
Words	Easy	2	18.19	
Words	Hard	1	16.56	
Words	Hard	2	18.31	1.75
Discourse	Easy	1	11.88	
Discourse	Easy	2	17.81	5.93
Discourse	Hard	1	15.25	
Discourse	Hard	2	10.56	-4.69

In spite of the agreement of two readability formulas, it is clear that the passages cannot be considered alternate forms. This effect points out the limitations of readability formulas as predictors of performance or difficulty in these tasks.

DISCUSSION

For the random word grouping students could be said to do equally well for hard or easy passages on either replication. For the discourse grouping the easy passages did yield higher (but not significantly higher) average scores than the hard passages, but there was disagreement between replications at the same difficulty level. This underscores the limitations of the readability formulas used. Two sets of passages, measured as equal in difficulty, were found to be significantly different by this group of students.

In hard passages, children did significantly better at the random word level. Context appears to have clear but not always facilitative effect on word boundary performance, perhaps implying these children were at least attending to a unit level larger than word by word.

The equal difficulty rankings given to alternate passages on the two readability formulas implies equality of word length and difficulty as well as similarity of sentence length. For this reason the random word versions of the alternate forms would be similar to one another. In contrast, the discourse passages likely accent the influence of added variables such as interest, propositional difficulty, and student experiential factors.

Two conclusions seem warranted in the present study. First, context seems to have some sort of effect, albeit uneven, on first grade students' ability to detect word boundaries. Second, researchers should worry about factors other than comparable readability levels when attempting to find alternate passages designed to serve as replications. One is reminded of Clark's (1973) advice about generalizing to populations of linguistic elements as well as populations of subjects.

REFERENCES

Clark, H. A. The language-as-fixed-effect fallacy: A critique of statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior,* 1973, *12,* 335-359.

Goldsmith, J. *The relationship of syntactic unit size to word boundary decisions of children — grades 2, 4, and 6.* Paper presented at the National Reading Conference, Kansas City, December 1974.

Goldsmith, J. *Word boundaries and syntactic units in the reading process.* Unpublished doctoral dissertation, Rutgers University, 1975.

Klein, H. A. & Klein, G. A. *Studying the use of context for word identification decisions.* Paper presented at the meeting of the National Reading Conference, New Orleans, December 1972.

Klein, G. A. & Klein, H. A. Word identification as a function of contextual information. *American Journal of Psychology,* 1973, *86,* 399-406.

Klein, H. A., Klein, G. A. & Bertino, M. Utilization of context for word identification decisions in children. *Journal of Experimental Child Psychology,* 1974, *17,* 70-86.

Klein, H. A., Klein, G. A., & Doris, A. *Some developmental trends in context utilization.* Proceedings, Eighty-first Annual Convention, American Psychological Association, 1973, 77-78.

Klein, H. A., Klein, G. A. & Hildum, D. Articulatory and constituent phrases as facilitators of word identification decisions. *Journal of Experimental Psychology,* 1974.

Klein, H. A., Klein, G. A. & Vigoda, C. H. *The utilization of contextual information by high school students.* Mimeographed, Oakland University, 1972.

JEROME A. NILES
AILEEN GRUNDER
CAROL WIMMER
Virginia Polytechnic Institute and State University

THE EFFECTS OF GRADE LEVEL AND SCHOOL SETTING ON THE DEVELOPMENT OF SENSITIVITY TO ORTHOGRAPHIC STRUCTURE

Sensitivity to orthographic structure is gained through only one route, experience in reading (Smith, 1971). The question arises at what point in the learning to read process is the reader able to make use of orthographic structure to aid in processing information? Gibson and Levin (1975) after a review of studies on the topic, concluded that toward the end of second grade, at least some children are able to demonstrate sensitivity to orthographic structure.

More specifically, Rosinski and Wheeler (1972) found that first grade children performed at chance level when differentiating among three- four- five- and six-letter nonsense word pairs containing an orthographically predictable and unpredictable word. A retest of the first grade students in the study one year later showed no significant improvement on the task. Golinkoff (1974), in a modified replication of this study, also found that under her visual condition first graders performed only slightly better than chance. Second graders were able to begin to make use of the inherent orthographic structure as they performed at significantly better than chance on the task.

Under different conditions there is evidence that readers are able to demonstrate their knowledge of orthographic structure as early as first grade. Lott and Smith (1970) set up a task which compared the amount of visual information needed to recognize a letter in isolation as opposed to its occurrence in a three-letter word. While there were differences among grade levels one through four in the ability to employ orthographic structure, first graders were still able to recognize letters in words more easily than letters in isolation. Niles (1976) used a similar methodology which revealed that letters in three- and four-letter nonsense words of high sequential dependency required less visual information for recognition than letters in low sequentially dependent words for first, third, and fifth grade Ss.

At this point a specific definition of the developmental trends of sensitivity to orthographic structure seems unclear. If one acquires this skill through reading, it would seem intuitively logical that the skill develops with the onset of reading. However, the question of what designates the onset of reading is in itself a question which has a number of answers. It is the intent of this study to more fully examine the continuum of learning to read in terms of orthographic structure by looking at students in grades kindergarten through sixth.

METHOD

Two schools were selected, one representing a student population drawn from a rural environment and a second representing an affluent suburban environment. Two intact classrooms were randomly selected at seven, end-of-year grade levels, kindergarten through sixth grade. The rural school had three classes at each level and the suburban school had four classes. This selection procedure, which omitted students who were absent on the day of testing, yielded approximately 90 to 100 students per grade level, with class size similar across the two schools; the total sample was 662.

A list of 15 pairs of eight-letter, zero order (ijhbwstt) and fourth order (ateditol) approximations to English (Miller, Bruner & Postman, 1954) was presented to each student on a single sheet of paper. The students as a class were instructed to mark the set of letters in each pair which looked most like a real word. To avoid an order effect, the random pairings were arranged in six different orders for each administration.

RESULTS

An initial summary of the means and standard deviations by grade level (Table 1) suggested a developmental trend in sensitivity to orthographic structure in English through the fourth grade. A one-way analysis of variance by grade level formally verified the existence of a main effect for grade level ($p < .001$).

Table 1

Means and Standard Deviations of
Correct Responses by Grade Level

	K	1	2	3	4	5	6
M	7.59	9.45	11.77	13.202	13.98	14.15	14.43
SD	2.13	2.4	3.13	2.96	1.58	1.88	1.44

The Newman-Keuls test for significant mean differences (Winer, 1971) indicated significant differences among all means through the fourth grade ($p < .01$). Significant differences did not exist among the fourth, fifth, and sixth grades ($p > .05$), although the fifth and sixth grade means did differ significantly with all other means, grade three and below. To avoid distortion of further data analysis because of the asymtotic nature of the data from fourth grade on, the two-way analysis of variance, schools x grade level, was applied to grades K - 4.

In addition to upholding the grade level effect ($p < .01$) the two-way analysis of variance revealed a main effect for schools. There was no significant interaction between schools and grade. As can be seen in Table 2, substantial differences were noted between schools at each grade level except first. The Newman-Keuls test formally verified these mean differences at the kindergarten, second, third and fourth grade levels as significant ($p < .01$).

Table 2

Means and Standard Deviation of Correct Responses by Grade Level and School*

	Grade:	K	1	2	3	4
Rural →	*M*	7.0	9.47	11.11	12.41	13.64
	SD	1.93	2.09	3.09	3.29	1.76
Suburban →	*M*	8.33	9.43	12.46	13.95	14.39
	SD	2.15	2.76	3.06	2.41	1.23
Difference →	*M*	1.33	.04	1.30	1.54	.75

*Maximum Score = 15

DISCUSSION

The linear trend of the data is rather consistent in its increments through the fourth grade level, at which point it appears that the reader has reached a point of maximum sensitivity to orthographic structure, at least for this task. These findings generally support an age-related notion of a reader's acquisition of sensitivity to orthographic structure. However an exception is taken with the conclusion (Rosinski & Wheeler, 1972) that this sensitivity begins to emerge for some children towards the end of second grade. The ability of the first grade children to perform significantly better than chance, while the kindergarten children performed at chance level, would seem to clearly demonstrate not only the presence of orthographic knowledge but the ability to apply this knowledge.

Sensitivity to orthographic structure should not be an unexpected skill for even a beginning reader to acquire. As pointed out by Smith (1975), the reader (especially the beginning reader) is faced with certain limitations to his information processing ability. The reader adjusts to these limitations by utilizing the redundancy available in language. When faced with the task of learning to read, the reader automatically employs his skills of economy in processing information and abstracts those principles which will reduce the demanding visual requirements of the task. The evidence supporting this notion is somewhat masked by the apparent age-related nature of sensitivity to orthographic structure. Operationally the grade level only serves as an index for the amount of experience a reader has with print. While this generalization holds for groups, it obviously breaks down for individuals. There were instances of children at the kindergarten and first grade levels with perfect scores. These kindergarten children were already considered capable readers by their classroom teachers and had automatically acquired maximum efficiency in dealing with orthographic structure because the task, learning to read, demanded it.

This same notion, the relationship between the degree of sensitivity to orthographic structure and the amount of experience with text, may also help to explain the between school difference for the rural and suburban schools. A close examination of the in-school and out-of-school reading environments for quality and quantity of on-going reading among the students would likely be most revealing in explaining the differences. One might postulate that more reading was going on in the suburban environment at the various levels, thus the suburban students had more opportunities to abstract the orthographic structure. This explanation is somewhat analagous, although for different reasons, to an avid first grade reader who demonstrates more knowledge of orthographic structure than the reluctant third grade reader. A good deal more documentation of specific reading environments must be done before conclusions can be firmly established for the given school differences.

The main effect for schools and the individual differences among specific children across grade levels does demand careful consideration when examining the age-related nature of sensitivity to orthographic structure. It does seem clear from this study that the development of sensitivity to orthographic structure is reciprocally related to learning to read. That is, one is aided in learning to read by developing this knowledge and one develops this knowledge by reading.

REFERENCES

Gibson, E. & Levin, H. *The psychology of reading.* Cambridge, Mass.: MIT Press, 1975.

Golinkoff, R. *Children's discrimination of English spelling patterns with redundant auditory information.* Paper presented to American Educational Association, February, 1974.

Lott, D. & Smith F. Knowledge of intraword redundancy by beginning readers. *Psychonomic Science,* 1970, *19,* 343-44.

Miller, G. A., Bruner, J. S. & Postman, L. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology,* 1954, *50,* 129-139.

Niles, J. The use of featural and letter dependency information in word recognition by elementary school children. *Reading Research Quarterly,* 1975-76, *11,* 198-202.

Rosinski, R. & Wheeler, K. Children's use of orthographic structure in word structure in word discrimination. *Psychonomic Science,* 1972, *26,* 97-98.

Smith, F. *Comprehension and learning.* New York: Holt, Rinehart and Winston, 1975.

Smith, F. *Understanding reading: A psycholinguistic analysis of reading and learning to read.* New York: Holt, Rinehart and Winston, 1971.

Winer, B. *Statistical principles in experimental design.* New York: McGraw-Hill Book Company, 1971.

RICHARD L. ALLINGTON
State University of New York at Albany

DEVELOPMENTAL TRENDS IN THE DISCRIMINATION OF HIGH FREQUENCY WORDS

In conjunction with other research focusing on understanding the nature of the development of visual perceptual abilities in beginning reading the present study examined the developmental trends in childrens' ability to visually discriminate among high frequency words.

Of the many processes involved in learning to read probably none has generated as much research and opinion as the visual perceptual process. Various visual perceptual deficits have been presented as primary factors in reading difficulties (Anapolle, 1967; Cruickshank, 1972; Frostig, 1967; Lyle, 1969; Vernon, 1957). Recent evidence (Allington et al., 1976; Kinsbourne, 1976; Lahey & McNees, 1975; Stanley & Hall, 1973; Steinheiser & Guthrie, in press; Vellutino et al., 1972, 1975, in press), however, has seriously undermined major aspects of the visual perceptual deficit hypothesis. These data do not argue that perceptual deficits are non-existent but rather that such deficits are not necessarily a general etiological factor in reading disability per se. Though in one text it was recently argued that we now understand the nature of the development of visual perceptual abilities (Spache & Spache, 1973) others argue that past research has provided only limited definitive data (Goldberg & Schiffman, 1972; Lahey & Lefton, 1976).

Several deficiences are common in the available research. The common experimental method has been to select groups of good and poor readers limiting the findings since subjects are typically drawn from a narrow age/grade span. That is, studies of older disabled readers, for instance, should not necessarily identify the identical deficits, if any, as the same study with younger participants (Satz, Rardin & Ross, 1971; Ohnmacht & Weiss, 1976). In the same vein, subjects for the various groups are typically selected from a common school population which may limit the results because of common and perhaps unique instructional experiences. The data of Barr (1975) and Cohen (1975) provide strong evidence that instructional effects can be potent factors in reading research. Another factor is the stimuli employed. Much research has used artificial alphabets (e.g., Gibson et al., 1962), artificial words (e.g., Marchbanks & Levin, 1965; Rayner, 1976), or single letters or letter strings (e.g., Popp, 1964; Lahey & Lefton, 1976). While these stimuli may provide useful insights they may not accurately assess discrimination abilities employed in reading. A final factor limiting much of the research on visual perceptual abilities relates to the validity and reliability of the tasks employed as dependent measures. Far too many studies have used experimenter constructed tasks which do not necessarily seem to tap abilities required for, or related to, the development of reading proficiency (Allington, 1976; Hammill & Bartel, 1975). Similarly, reliability coefficients too often go unreported; some studies have used so few measures of a skill it seems unlikely that the reliability coefficient would have been reported had it been computed.

The present study sought to assess skills related to reading achievement as directly as possible; the stimuli were high frequency English words. Subjects were drawn from four grade levels and ten school districts in an attempt to ensure greater generalizability of the findings. The primary purpose of the study was to examine the developmental nature of visual discrimination abilities. An earlier study (Allington, in press) had provided data indicating that mid-year first graders averaged better than 90% accuracy on visual discrimination tasks involving high frequency words, but since subjects from only a single grade were employed no judgments about later proficiency levels could be reported. Recently it was suggested that untimed visual discrimination deficiencies may play a role in reading difficulties at the elementary level (Lahey & Lefton, 1976). However, their experimental task employed consonant strings rather than words and while poor readers at grades 2 and 3 were found deficient when compared to good readers on this ability no suggestion as to an acceptable level of proficiency was made. The present study not only attempts to examine developmental trends in visual discrimination of words but also attempts to roughly define levels of normal visual perceptual abilities. Additionally, error patterns are examined to establish whether commonality exists across age/grade levels.

METHOD

Subjects

Two hundred seventy five subjects participated in this study. Approximately equal numbers of boys and girls were drawn from grades K - 3 in 10 school districts identified as urban, suburban, or rural communities.

All testing was carried out in the fifth of six months of the school year. Since all participating schools randomly assigned children to classrooms the cooperation of a single teacher at a selected grade was elicited. Twenty-five subjects were randomly identified in classes where enrollments exceeded this number and while all students in a given classroom were administered the experimental materials the data from only the 25 randomly selected subjects were subjected to analysis.

Materials

The stimuli employed were the total number of words ($N = 120$) which appear on each of three widely used basic word lists; Dolch (1920), Durr (1973), and Harris-Jacobsen (1972) with the requirement that only words at or below the core first reader level from the Harris Jacobson list be selected.

The target words were embedded among four distractor words. For each item, the four distractor words were selected adhering to the following constraints: one word was drawn from the word lists which shared at least two letters for words of four or more characters or one letter for words less than four characters in length; a second word also following the above constraints but not necessarily drawn from the word list; a third word which began with the same letter and had the same number of letters as the target word and a fourth word which began with the same letter but had a different number of letters than the target word. Within these constraints words were sought which would maximize the discriminability difficulties for each item.

After the distractor words for each item were selected one was randomly assigned to the first position in the item since prior field testing indicated virtually no errors occurred when the correct word choice was in position 1, or immediately to the right of the target word. The remaining distractors and the correct word choice were then randomly assigned to the remaining four positions, as in examples (1) and (2)

(1) come | home come came can cone

(2) this | with his thin this thing

The stimuli were presented to the students on five separate worksheets which were administered singly over five separate instructional days. Each worksheet presented 24 target words printed in primary type, all lower case letters. The worksheets presented black stimuli on white 8-1/2" x 11" unlined paper. The subjects were instructed to "mark the word that looks exactly like the one behind the line." Practice examples were presented in an attempt to ensure each subject understood the nature of the task.

Teachers participating in the study were told not to correct the students' papers; the student responses were transferred to IBM coding sheets for computer scoring and item analyses.

RESULTS

All scores used in the analysis were the sum of errors and omitted items subtracted from the total possible score of 120.[1] The mean scores and standard deviations for all groups are presented in Table 1.[2] Even a brief perusal of these data show the developmental trend for improvement on this visual discrimination task. However, while tremendous gains are evident between kindergarten and grades 1 and a lesser improvement between grades 1 and 2, there is a leveling effect at grade two with virtually no gain between grades 2 and 3. At grade 2 the mean equalled a 97% accuracy level. Given the nature and length (5 consecutive days) of the experimental procedures it would be surprising if much improvement could be made in this accuracy level.

[1] Some data were reconstructed since certain subjects, as described below. were absent and were not able to later complete one of the experimental tasks (sheets). For these missing data the mean of the subjects group (school type, sex, and grade) on the missing sheet was substituted for the missing score. Fortunately, performance on single sheet was not available for only 13 of the 275 subjects as follows: 1 female and 2 male suburban kindergartens, 2 female and 3 male urban first graders, 2 male rural kindergartens, 1 female and 1 male rural first graders, and 1 rural second grader. Educators at the urban schools contacted declined to administer the task to kindergarten subjects arguing it would be too difficult.

[2] Some kindergarten subjects were unable to do the experimental task as evidenced by 1) complete refusal to engage in task, or 2) an inordinate amount of time (70 minutes) to complete 5 items on the first sheet (incorrectly). Two male suburban subjects, 1 female and 2 male rural subjects were unable to do the task. Their scores of 0 were used in the analysis.

Table 1

Means and Standard Deviations
Number of Correct Matches to Standard

School Type		Urban		Suburban		Rural	
Grade		Male	Female	Male	Female	Male	Female
Kindergarten	*M*			86.6	100.83	68.2	93.3
	SD			29.84	25.49	28.20	23.63
1st	*M*	107.2	108.8	110.1	107.1	105.1	109.0
	SD	3.15	7.35	5.88	15.14	20.08	8.96
2nd	*M*	114.5	117.6	115.6	118.0	116.85	117.1
	SD	12.44	2.65	4.08	1.71	4.20	2.88
3rd	*M*	116.4	115.3	118.1	118.2	118.5	117.8
	SD	4.09	4.5	1.83	2.27	1.56	1.78

The reliability of the experimental tasks were computed using two common techniques. The reliability coefficient using the split-halves technique was .96. A similar coefficient ($r = .95$) resulted when the Kuder-Richardson technique was employed. Both coefficients approximate the reliability level obtained in the earlier study (Allington, in press).

An analysis of variance was applied to the data for only the 225 subjects in grades 1, 2, and 3. This was necessary since there were not data for urban kindergarten subjects thus creating an empty cell in a 4 (grade level) x 3 (school type) x 2 (sex) design. Thus, the ANOVA was performed on the three grade levels (1, 2, 3) with intact data for all school types. Only the effect for grade was significant, $F(2,207) = 32.873, p < .01$. A Scheffé test indicated that grade 1 subjects' performance was significantly different ($p < .01$) from the performances of grade 2 and grade 3 subjects but the performance of grade 2 and grade 3 subjects did not differ.

DISCUSSION

An important finding was the lack of effect for either the school type (urban, suburban, rural) or the sex variables in grades 1, 2, and 3. This contradicts a widespread myth that these variables play a significant role in the development of the prerequisite skill assessed here: visual discrimination of words. Unfortunately the incomplete data for the kindergarten subjects (see Note 1) precludes an attempt to compare subjects at this earlier level. In any event, if differences should exist across these variables at kindergarten they are evidently overcome by the middle of first grade.

The significant effect for grade is not surprising nor are the lack of differences between grades 2 and 3. These data confirm the hypothesis that visual discrimination of frequent words is well developed (near 90%) by the middle of first grade and a ceiling effect was experienced at the middle of grade 2. While subjects were not identified by reading ability level it seems unlikely that visual discrimination difficulties play any significant role in reading disability beyond grade 1. A similar conclusion was recently drawn by Satz et al. (1975) who reported that the significant differences found between two groups of kindergarten subjects on recognition-discrimination tasks had disappeared when these subjects were retested in grade 2, even though the group which had performed poorly earlier were now severely disabled readers. Thus it seems that while visual discrimination at prereading assessments may be an effective predictor of later reading success, beyond kindergarten, or early grade 1, performance of poor readers will equal that of good readers (cf. Allington et al., 1976; Kinsbourne, 1976; Lahey & McNees, 1975; Vellutino et al., 1972). The misreading of words is too often construed as indicative of discrimination deficiencies without adequate support for such a conclusion. The data gathered to date suggest the etiology of the misreading will typically lie elsewhere, at least at grade 2 and beyond. Thus programs, or instruction, designed to correct visual perceptual deficiences would not seem constructive beyond that level.

The data presented in this paper may also be useful for identifying what Kinsbourne (1976) has called "external referent points." He states we currently have little information about criterial levels for many processes basic to reading achievement. Without these referents we may seek efficiency levels which are redundant in that they exceed the level necessary to achieve the task. (A recent paper [Walmsley, 1975] demonstrates the validity of this assumption for tests of trigram recognition, often used to assess phonic skills.) The data in this study may serve as useful guidelines for predicting criterial levels on similar tasks. Unfortunately, the study was not designed to provide answers to this issue and thus the data are less appropriate than might have been.

REFERENCES

Allington, R. L. A note on the Jordan Left-Right Reversal Test. *Academic Therapy*, 1976, *11*, 409-414.

Allington, R. L. The visual confuseability of high frequency words. *Journal of Learning Disabilities*, in press.

Allington, R. L., Gormley, K. & Truex, S. Poor and normal readers achievement on visual tasks involving high frequency low discriminability words. *Journal of Learning Disabilities*, 1976, 9, 292-296.

Anapolle, L. Visual training and reading performance. *Journal of Reading*, 1967, *10*, 372-382.

Barr, R. The effect of instruction on pupil reading strategies. *Reading Research Quarterly*, 1973, *10*, 555-582.

Cohen, A. S. Oral reading errors of first grade children taught by a code emphasis approach. *Reading Research Quarterly*, 1975, *10*, 616-650.

Cruickshank, W. M. Some issues facing the field of learning disability. *Journal of Learning Disability*, 1972, *5*, 380-383.

Dolch, E. W. A basic sight vocabulary. *Elementary School Journal*, 1936, *36*, 456-460.

Durr, W. K. A computer study of high frequency words in popular trade journals. *Reading Teacher*, 1973, *27*, 37-42.

Frostig, M. Education of children with learning disabilities. In E. C. Frierson & W. B. Barke (Eds.), *Educating children with learning disabilities*. New York: Appleton, Century, Crofts, 1967.

Gibson, E. J.; Gibson, J. J.; Pick, A. D.; & Osser, H. A developmental study of discrimination of letter-like forms. *Journal of Comparative and Physiological Psychology*, 1962, *55*, 897-906.

Goldberg, H. K. & Schiffman, G. *Dyslexia: Problems of reading disabilities*. New York: Grune and Stratton, 1972.

Hammill, D. D. & Bartel, N. R. *Teaching children with learning and behavior problems*. Boston: Allyn-Bacon, 1975.

Harris, A. J. & Jacobsen, M. *Basic elementary reading vocabularies*. New York: MacMillan, 1972.

Kinsbourne, M. Looking and listening strategies in beginning reading. In J. T. Guthrie (Ed.), *Aspects of reading acquisition*. Baltimore: Johns Hopkins, 1976.

Lahey, B. B. & Lefton, L. A. Discrimination of letter combinations in good and poor readers. *Journal of Special Education*, 1976, *10*, 207-210.

Lahey, B. B. & McNees, M. P. Letter discrimination errors in kindergarten through third grade: Assessment and operant training. *Journal of Special Education*, 1975, 9, 191-199.

Lyle, J. G. Reading retardation and reversal tendency. *Child Development*, 1969, *40*, 833-843.

Marchbanks, G. & Levin, H. Cues by which children recognize words. *Journal of Educational Psychology*, 1965, *56*, 57-61.

Ohnmacht, F. & Weiss, F. Construct validity and reading research: Some comments with a data simulation exemplar. *Journal of Reading Behavior*, 1976, *8*, 19-25.

Popp, H. M. Visual discrimination of alphabet letters. *Reading Teacher*, 1964, *17*, 221-226.

Rayner, K. Developmental changes in word recognition strategies. *Journal of Educational Psychology*, 1976, *68*, 323-329.

Satz, P.; Rardin, D.; & Ross, J. An evaluation of a theory of specific developmental dyslexia. *Child Development*, 1971, *42*, 2009-2021.

Satz, P.; Friel, J.; & Goebel, R. A. Some predictive antecedents of specific reading disability: A three-year follow-up. *Bulletin of the Orton Society*, 1975, *25*, 91-110.

Spache, G. D. & Spache, E. *Reading in the elementary school*, (3rd ed.). Boston: Allyn Bacon, 1973.

Stanley, G. & Hall, R. A comparison of dyslexic and normals for recalling letter arrays after brief presentation. *British Journal of Psychology*, 1973, *43*, 301-304.

Steinheiser, R. & Guthrie, J. T. Reading ability and efficiency of graphemic-phonemic encoding. *Quarterly Journal of Experimental Psychology*, in press.

Vellutino, F. R.; Steger, J. A.; Kaman, M.; & Desetto, L. Visual form perception in deficient and normal readers as a function of age and orthographic-linguistic familiarity. *Cortex*, 1975, *11*, 22-30.

Vellutino, F. R.; Steger, J. A. & Kandel, G. Reading disability: An investigation of the perceptual deficit hypothesis. *Cortex*, 1972, *8*, 106-118.

Vellutino, F. R.; Steger, B. M.; Moyer, S. C.; Harding, C. J. & Niles, J. A. Has the perceptual deficit hypothesis led us astray? *Journal of Learning Disabilities*, in press.

Vernon, M. D. *Backwardness in reading*. Cambridge: Cambridge University Press, 1957.

Walmsley, S. A. *The criterion-referenced measurement of early reading behavior*. Paper presented at the Annual Meeting of the New England Educational Research Organization, Provincetown, Mass., May 1975.

CHARLES READ
University of Wisconsin, Madison

WHAT LEVEL OF REPRESENTATION DO CHILDREN EXPECT TO FIND IN PRINT?

LEVELS OF REPRESENTATION

In the analysis of language, one can distinguish several levels, from that of the sounds themselves through the phonemic, the morphemic, and the word level, on to the level of sentences, with various levels of syntactic representation, up to the level of meanings. In what follows, "abstractness" will mean increasing distance from the level of speech sounds and increasing closeness to the level of meaning.

These levels of linguistic analysis pertain to the present question, for there is also a hierarchy of levels on which language can be represented in writing, and the two hierarchies correspond at some important points. Table 1 lists some general types of writing systems against the levels of language which each represents.

Table 1

Types of Writing Systems

Writing System	Level of Language Represented
Pictographic	Meanings, or messages, without necessarily representing individual words
Logographic	Words, without necessarily representing components of words such as syllables or segments
Syllabic	Syllables, without necessarily representing individual segments
Alphabetic	Segments, of which two degrees of detail may be distinguished:
Phonemic	Distinguishes only those segments which are "significantly" distinct
Phonetic	Distinguishes, at least in principle, all sounds which can be reliably recognized as different

This table is akin to, but partly distinct from, classic taxonomies of writing systems, such as that in Gelb (1963, figs. 2, 3, and 95).

An instance of a pictograph might be a representation of a canoe and a paddler, where the intent is not to *depict* a canoe and a paddler, but rather to convey a message such as "this point is on a scenic river trail," or "canoes for rent." Note that the words in the message need not correspond to elements in the picture. A familiar example is the set of International Road Signs, such as the one that means "this is a school crossing," on which there are two stick figures with rectangular objects in their hands, walking between parallel lines. Nothing directly represents the word "school."

An example of logographic writing is, of course, the Chinese system, with its 50,000 logographs whose parts correspond to units of meaning rather than to units of sound. For syllabic writing, Japanese furnishes a familiar example, although the syllabaries exist side-by-side with logographic characters in Japan. Note that syllabaries do not necessarily encode information about individual segments: for example, in the Japanese katakana and hiragana syllabaries, the symbols for /ka/, /ki/, and /ku/ have no elements in common (Gelb, 1963, fig. 86).

Among the alphabetic systems, which do represent sound segments, we must distinguish the phonemic from the phonetic, the latter being those which make it possible to distinguish any two sounds which a trained listener can reliably recognize as different in phonetic quality. Thus, a phonetic representation can distinguish the two /k/ sounds in "keep cool," since they are pronounced slightly differently, whereas a phonemic representation would treat them alike, since they are functionally equivalent in English, that is, they never serve to distinguish one word from another. Fries (1963, ch. 5) provides a reliable account of this sometimes-abused distinction.

Even though a phonetic representation seems to bring us close to the speech sounds themselves, closer, indeed, than most people have any real desire to be, it is worth noting that a phonetic representation is still an abstraction with respect to the actual sounds, the physical events. A phonetic representation still involves categorization, putting sounds which may be infinitely varied into a finite number of types, and it involves segmentation, cutting the continuous stream of vibrations in the air into discrete units. By contrast, a sound spectrogram preserves the acoustic event on paper in immense detail; it does not categorize (at least not in the way that humans do) and it does not segment. Speech appears as more-or-less continuously varying shadings, the darkness of which corresponds to the intensity of vibrations at various frequencies. Precisely *because* it is too faithful to the physical signal, because it is not abstract, a spectrogram cannot be read, in the usual sense. For all practical purposes, people cannot learn to decode the message from a spectrogram reliably; it is not a writing system. Perhaps surprisingly, then, not only do writing systems vary in their abstractness, but they must be somewhat abstract in order to be useful for the ordinary purposes.

With this brief survey of the types of writing systems and their degrees of abstractness from the level of sounds, we can now approach the question, "what level of representation are children best prepared to find in orthography as they first approach reading and writing?" In passing, let us take up three related questions:

— What type is best for purposes other than pedagogical?

— What has been the historical development of these writing systems?

— Where does English orthography fit in this hierarchy?

On all three questions, we have made assumptions which are inadequate or at least not yet substantiated.

First, as to selecting the best orthography for non-pedagogical purposes, the beginning of wisdom is to recognize that the choice depends significantly on the purposes. We do in fact use pictograms, as in the international road signs; for rapid communication of a message, independent of the reader's native language, where the precise wording is not important, pictograms seem to have no equal. In our increasingly multilingual and rapidly moving society, it is quite possible that they will find more uses. It sometimes seems that the more urgent the message, the more likely it is to be conveyed pictographically; some guides and warnings become pictograms, in effect, because of their rigidly conventionalized presentation. One doesn't have to know English in order to identify the exit signs in a theater or an airplane.

Also, we find good uses for logograms. Numerals, the ampersand, the plus sign, and even that potent symbol, the dollar sign, are all logograms whose usefulness is hardly declining. It would be extraordinarily cumbersome to catalog automobile parts or to do arithmetic with numbers written out. However strongly we may believe that the invention of alphabetic writing marked the beginning of Western civilization, the fact is that other systems are not mere relics of antiquity but are appropriate and indeed superior for some purposes.

Ordinary English orthography, too, finds a great deal of use for representations which are as much logographic as they are alphabetic. The written distinctions between homophones such as *buy* and *by, see* and *sea, steel* and *steal* can be regarded as adaptations of our alphabetic writing which now serve to maintain distinctions between words and morphemes, even though many of the spellings originated from differences in pronunciation. Even more pervasive is spelling a given morpheme consistently, despite variations in pronunciations. Thus the past tense of regular verbs is spelled *-ed,* even though it receives one of three different pronunciations, depending on the preceding segment. Similarly, the same meaning unit in *wid*e and *wid*th, *sign* and *sign*al, *please* and *pleas*ant, *part* and *part*ial receives a consistent spelling despite the differences in pronunciation. In this respect English orthography is neither precisely logographic nor alphabetic, but clearly compromises in the direction of logography by representing meaning relationships instead of representing phonetic segments. This characteristic of English orthography has been called "morphemic" (Venezky, 1970, p. 120) and "lexical" (C. Chomsky, 1970). These same writers and several others have suggested that it may well be an aid to the adult reading silently for meaning, lending further support to the argument that no one level of representation is optimal for all purposes. For the reader who doesn't know what "logography" means (and it is not found in *Webster's Third New International Dictionary* (1971)), it seems clearly advantageous for *logo-* and *-graph-* to be spelled in their consistent ways, rather than to represent differences in pronunciation between [logógræfɪk] and [loˈgɔgrəfi], the pronunciations found in the *American Heritage Dictionary of the English Language* (1969). Linguists and non-linguists alike have assumed that a phonemic representation is optimal for both reader and writer, adult and child, but this assumption is increasingly coming into question.

In the light of these observations, there is an historical belief which needs to be revised. Many of our introductory textbooks suggest that the march of orthography has been inexorably toward alphabetic writing. J. N. Hook (1975, pp. 40-41), for example, opines that if the Chinese had not been burdened with their logographic writing system, they might well be masters of the world by now. Even so reputable a scholar as Gelb has a section entitled "Alphabet's Conquest of the World," although he by no means disparages non-alphabetic writing systems. There is indeed some support for these notions; even the People's Republic of China has begun a gradual conversion to alphabetic writing ("PRC converts to Romanized alphabet," 1976). Nonetheless, the situation is clearly more complex. Every culture finds some use for logographic and pictographic representation. English orthography has actually moved toward greater abstractness; before the Early Modern period, it was more consistently phonemic than it is now. True enough, some of the change resulted from 18th and 19th-century neoclassicism, and we attribute much of the rest to the conservatism of printing and schooling; as far as we know, orthographic change in English did not arise in deliberate response to the needs of the reader. Much of the change has not been in orthography at all, but in pronunciation, while the spelling changed much more slowly. In that sense one might say that we backed into greater abstractness, rather than embracing it. Still, it is at least arguable that when it represents meaning relationships, the abstractness of our orthography is an aid to the reader, and that that quality helps to account for its resistance to change. Across the millenia there has indeed been a development from pictographic to alphabetic writing, but there are surely limitations and constraints on this process. We are not entitled to conclude that the more concrete an orthography is, the better it is for all its users.

CHILDREN

At last, we come to the key question: which type(s) of representations are children best prepared to encounter? In using a phrase like "best prepared" or even more metaphorically, "expect to find," I do not mean to suggest that children are aware of levels of abstraction, nor do I assume that there must be a single level which is best for all children at all times. My aim has simply been to sketch a background against which this question can be considered without prejudice, to indicate why it is not absurd to ask whether children might actually learn most easily to read representations that are not precisely phonemic, and maybe not even alphabetic.

The evidence is far from satisfactory. I mean to be posing a research question, rather than answering one, and it is not an easy question to investigate. Let me begin with the anecdotal evidence, which I find most stimulating, if not necessarily most convincing. There is some evidence on each side. For example, another scholar once mentioned that her daughter wanted to devise two different spellings for the two /k/ sounds of *cocoa,* in order to represent a slight phonetic difference between them. That is phonetic spelling indeed. But in my experience the more typical examples reveal a readiness for abstraction. In observing my own son learning to read, I have been struck, for example, by how easily he absorbed the fact that *-tion* is a regular ending whose pronunciation is reliable, even though it is not directly indicated by the spelling.

Far more general is the observation that in their invented spelling, even children who otherwise rely on letter-names in making up spellings quickly adopt the standard use of *s* to spell the sound /z/. The effect is a more abstract spelling of the plural and possessive endings and the third-person present-tense verb inflection. For example, one child produced the spelling FES for the word *friends*. Comparing this not to the standard spelling but to the phonetic form [frɛnz], it is obviously a very immature spelling,

F E S

[frɛnz]

The /r/ of the initial cluster is omitted, along with the preconsonantal nasal, the latter b[illegible] a very frequent omission in young children's spelling, even through the first grade. But amidst this immaturity, there is one strikingly standard spelling: the plural is spelled *s* even though it is pronounced [z]. I have presented this example as an anecdote, but it is actually very representative. In my collection of children's invented spellings, the plural, possessive, or present tense [z] ending is spelled *z* only twice in 150 examples. From Berko's research (1958), we know that children can pronounce plural endings correctly in the early grades, but the spellings suggest that at another level, children are prepared to represent the plural as a single unit, despite its different pronunciations. I believe that this spelling has its basis in children's sense of the similarity between [s] and [z] (Read, 1975, ch. 6) as well as the influence of standard spelling, but the effect is to produce a more abstract representation of the plural and possessive endings.

Directly pertinent to our hierarchy of representations is the work of Lila Gleitman and Paul Rozin in teaching American children to read a syllabary. Rozin, Poritsky, and Sotsky (1971) reported that "American children with reading problems can easily learn to read English represented by Chinese characters." Appearing in *Science* magazine, this brief account caught the fancy of many scholars, within and outside the field of reading. Gleitman and Rozin (1973) followed up with a more extended project. Gleitman's conclusion, not to put it too summarily, was that she had won the battle but not the war: children with "a poor prognosis" for success in reading learned to read a syllabary but still had the expected amount of difficulty in learning to read English written alphabetically. Similarly, I. Y. Liberman, et al. (1977) have reported several pieces of evidence that phonetic segmentation, dividing an utterance into the correct number of segments, is a formidable task for some beginning readers and a task which discriminates between good and poor readers.

Taken together, this evidence suggests to me that children may be in some respects better-prepared to encounter a syllabic or a logographic representation than an alphabetic one in their approach to reading and writing. We have too readily assumed that it is the abstractness of English orthography, rather than the concreteness of it, which leads to difficulty in learning to read. However, it seems that access to segmentation and the use of alphabetic correspondences is essential not only to initial reading but to skilled performance (Rozin and Gleitman, 1977; Venezky and Massaro, in press).

CONCLUSIONS

I conclude, then, that three common assumptions have been too simple:

In comparing orthographies, we have failed to recognize that representations at the syllable, word, and even meaning level have their uses and that there is no such thing as an orthography which is optimal for every purpose.

In teaching about the history of orthography and the general tendency toward alphabetic writing, we have failed to note that English orthography has actually moved toward greater abstractness in the last four centuries.

In observing children we have failed to take sufficient notice of the evidence that they may be prepared for certain abstractions in writing. We have looked for irregularity and consequent difficulty where English spelling fails to be phonemic, without adequately considering whether regularity at other levels may be accessible to children. English spelling assuredly presents difficulties, but phonemic regularity may not be the best or the sole criterion.

Turning simultaneously to pedagogy and research, I wish to propose a small experiment with the initial teaching alphabet (i.t.a.). I.t.a. uses three characters to present /s/ and /z/: *s*, *z*, and ƨ. The backwards *z* or angular *s* represents /z/ sounds which are spelled *s* in standard orthography, not only in plurals, possessives, and verb endings, but also in words like *was*, *is*, *his*, *always*, *resume*, etc. If my observations are correct, most children at the first-grade level and beyond are quite prepared to ignore this distinction between /s/ and /z/, and the third character is unnecessary.

I predict that eliminating this character would facilitate both the use of i.t.a. and the transition to standard orthography. Indeed, I would make the same prediction about the two characters ʈh and ŧh, which mark the distinction between /θ/ and /ð/, as in *thin* and *then*, *thigh* and *thy*. This change would make i.t.a. somewhat less true to its own principles, but it contains many compromises already. In making this proposal, I am not claiming that i.t.a. is either good or bad, but that it is not sufficiently abstract in these particular cases.

This small proposal is just one example of an area of research which I feel we can and should strengthen: the question of what levels of representation are most accessible to children in the preschool and primary years. I am quite prepared for a great deal of individual variation, but I am also prepared to find that children are capable of categorizing and abstracting above the phonemic level.

REFERENCES

American heritage dictionary of the English language. New York: American Heritage Publishing Co., 1969.

Berko, J. The child's learning of English morphology. *Word*, 1958, *14*, 150-177.

Chomsky, C. Reading, writing, and phonology. *Harvard Educational Review*, 1970, *40*, 287-309.

Fries, C. C. *Linguistics and reading*. New York: Holt, Rinehart and Winston, 1963.

Gelb, J. I. *A study of writing* (2nd ed.). Chicago: University of Chicago Press, 1963.

Gleitman, L. R. & Rozin, P. Teaching reading by use of a syllabary. *Reading Research Quarterly*, 1973, *8*, 447-483.

Hook, J. N. *History of the English language*. New York: Ronald Press, 1975.

Liberman, I. Y., Shankweiler, D., Liberman, A. M., Fowler, C., & Fischer, F. W. Phonetic segmentation and recoding in the beginning reader. In A. S. Reber & D. Scarborough (Eds.), *Toward a psychology of reading*. New York: Halsted Press, 1977.

PRC converts to Romanized alphabet. *The Linguistic Reporter*, 1976, *19*, 3.

Read, C. *Children's categorization of speech sounds in English*. Urbana: National Council of Teachers of English, 1975.

Rozin, P. & Gleitman, L. R. The structure and acquisition of reading. In A. S. Reber and D. Scarborough (Eds.), *Toward a psychology of reading*. New York: Halsted Press, 1977.

Rozin, P., Poritsky, S., & Sotsky, R. American children with reading problems can easily learn to read English represented Chinese characters. *Science*, 1971, *171*, 1264-1267.

Venezky, R. L. *The structure of English orthography*. The Hague: Mouton, 1970.

Venezky, R. L. & Massaro, D. W. The role of orthographic regularity in word recognition. In L. Resnick & P. Weaver (Eds.), *Theory and practice of early reading*. Hillsdale, N.J.: Erlbaum Press, in press.

Webster's third new international dictionary of the English language, unabridged. Springfield, Mass.: G. & C. Merriam Co., 1971.

JOSEPH L. VAUGHAN, JR.
University of Arizona

ROBERT J. TIERNEY
University of Arizona

MADELON ALPERT
Sunnyside High School

A PSYCHOLINGUISTIC ANALYSIS OF CLOZE RESPONSES

Since the advent of tranformational linguistic theory (Chomsky, 1957), increased efforts have been made to establish a better perspective of reading comprehension. More precisely, various efforts have been made to examine responses to written language in terms of cueing systems (Goodman, 1965; Weber, 1968). Goodman (1969) has proposed and used an in-depth analysis of children's oral responses to a reading selection for purposes of (1) evaluating pupils' use of grapho-phonemic, syntactic, and semantic cueing systems and (2) helping to formulate a psycholinguistic model of the reading process. As research efforts are undertaken to examine reading comprehension from a psycholinguistic perspective, it would seem that basic to such efforts is the question of the relationship between syntactic and semantic cueing systems, especially in light of the existent disagreement among various linguists including Chomsky (1957, 1965), Lakoff (1972), and Rommetveit (1974). This study was designed to examine the nature of the relationship between syntactic and semantic cueing systems in reading comprehension. More specifically, the purpose was to examine the relationship between syntactically acceptable and semantically acceptable responses by various pupil groups on cloze tasks.

METHOD

Two hundred forty students on four grade levels (four, six, eight, and eleven) were asked to read two selections and respond to cloze tests immediately after reading each selection. On each grade level, the students were identified as good, average, or poor readers on the basis of their performance on a standardized reading test. The selections they read were determined to have a readability level commensurate with the students' grade level. The Fry readability graph was used for this purpose. To incorporate the effect of interest on reading comprehension, each student selected one passage of high interest and one passage of low interest from brief descriptions of seven selections on their own grade levels.

Students were asked to read each selection and no time limit was imposed. As they finished reading, each student was directed to fill in every blank on a cloze test, i.e., "leave no blank empty," as this was considered especially important for assessing syntactic perceptions. Thus, even if a reader could not ascertain semantic sense for a blank, the response could be examined for syntactic acceptability.

Each cloze test consisted of fifty deletions using an every fifth word deletion pattern. An analysis of each response was made in accordance with semantic and syntactic appropriateness for each deletion. Semantic acceptability was based on three categorical definitions: (1) exact replacement, (2) synonym, or (3) minimal change. The exact response category required the proper term as used by the author with misspelled words being accepted if the intended response was clear. For a response to be an acceptable synonym, the meaning was required to remain unchanged and the word used was required to agree in number, gender, and tense with the deleted item. The category of minimal change allowed for semantically acceptable responses where, through connotation for example, a slight shift of meaning might occur. The sum of these three semantic categories was computed to determine each student's semantically acceptable score. In the analysis of the syntactic-semantic relationships, the summed score was the one used to represent the students' semantic awarenesses. Syntactic acceptability of a response was determined by agreement in form class (part of speech) *and* grammatical function with the deleted item. This score, when determined to be independent of the semantic score, was used as the index of the students' syntactic perceptions. Interjudge reliability was checked periodically to ensure agreement among the three assessors, and these checks indicated approximately 90 per cent agreement.

RESULTS

Syntactic and semantic scores were determined for each student on each of the two passages. The relationship between the total semantic score and the syntactic score was examined by Pearson product-moment correlations on all grade levels. The results indicated very high positive correlations (.79 to .89) on all grade levels with both high and low interest material. One would expect such findings due to the function of the scoring system used to determine the semantic and syntactic scores. Rarely will a response on a cloze test be unacceptable syntactically when it is semantically acceptable. Thus, the syntactic score is confounded by the semantic score due to an artifact of the scoring system. If correlations between these factors were to suggest the true construct relationships, the artifact of the scoring system had to be eliminated.

A new syntactic score was derived which did *not* include the semantically acceptable score. This new syntactic score was the proportion of non-semantically acceptable responses which were syntactically acceptable. (For example, assume that, of fifty responses, thirty were semantically acceptable. Each of the remaining twenty responses was examined to determine if it was syntactically acceptable. If ten were syntactically acceptable, the syntactic score was fifty per cent.) Among the responses scored in this study, the mean scores of syntactically acceptable responses which were *not* semantically acceptable were, for good, average, and poor readers, 11.245, 14.26, and 11.80, respectively. Unfortunately, the limitation of the syntactic element within the semantic score remains, and as long as the semantic factor is to be examined within a performance framework, it is unlikely that this can be avoided.

The relationship between the "new" syntactic score and the semantic score was examined by Pearson product-moment correlations. These were computed for each grade level. The results (Table One) indicated moderately high positive correlations on both the high interest and low interest passages. When corrected for attenuation, the correlations increased slightly. This moderately positive correlation between syntactically acceptable and semantically acceptable scores was also found within each ability group on all four grade levels.

Table 1
Correlations of Syntactically Acceptable and
Semantically Acceptable Cloze Responses

		High Interest		Low Interest	
Grade	N	*r*	*r* corrected for attenuation	*r*	*r* corrected for attenuation
4	62	.49	.60	.70	.88
6	64	.61	.76	.58	.72
8	58	.52	.64	.66	.82
11	56	.63	.78	.63	.78

$p <$ .001 for all correlations

CONCLUSION

The moderately high correlations between syntactically acceptable and semantically acceptable cloze responses suggest that, as common sense infers, semantic and syntactic elements of language are neither independent of one another nor totally dependent. They *are* highly related elements of language, and when one seeks to explain language and the relationship of students and reading achievement, both of these elements should be considered. In terms of assessing observed responses in reading comprehension via cloze tests, a meaningful relationship exists between syntactic and semantic elements. Although Chomsky (1957) has maintained that syntax and semantics are independent aspects of language, Rommetveit (1974), Lakoff (1972), as well as other generative grammarians disagree. As related to cloze responses, the findings here tend to support the latter perspective and suggest that additional empirical research is needed to provide further insights into the true construct relationships.

REFERENCES

Chomsky, N. *Syntactic structures.* The Hague: Mouton Press, 1957.

Goodman, K. A linguistic study of cues and miscues in reading. *Elementary English,* 1965, *42,* 39044.

Goodman, K. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly,* 1969, *5,* 9-30.

Lakoff, G. Linguistics and natural logic. In D. Davidson and G. Harman (Eds.), *Semantics of natural language.* Dordrecht: Reidel, 1972, 545-665.

Language and Mind. New York: Harcourt, Brace, and World, 1968.

Rommetveit, R. *On message structure.* Bristol: John Wiley and Sons, 1974.

Weber, R. The study of oral reading errors. *Reading Research Quarterly,* 1968, *4,* 96-119.

R. SCOTT BALDWIN
The University of Tulsa

CLAUSE STRATEGIES AS A FACTOR IN READING COMPREHENSION

According to Goodman (1976),

> In silent reading, the reader sweeps ahead, sampling from the graphic input, predicting structures, leaping to quick conclusions about the meaning and only slowing down or regressing when subsequent sampling fails to confirm what he expects to find (p. 482).

This statement captures the essence of a psycholinguistic theory as it exists in the field of reading today; and there is something intuitively appealing about viewing the reader as one who predicts grammatical structure, tests linguistic hypotheses, and extracts and analyzes minimal language cues. Yet there exists a basic set of mysteries which we, as psycholinguistic sleuths, have yet to come to grips with. For example, why are some linguistic cues selected from the perceptual input while others are ignored? What precisely are the cognitive conditions and the textual circumstances which render some language cues salient and others incidental? And what are the "specific" strategies which readers employ in predicting grammatical structures?

The author's perception is that we in the field of reading are fond of shifting the explanatory burden for psycholinguistic phenomena to the linguists and the experimental psychologists. We are inclined to think that if a child comes to school with a reasonably well established linguistic competence his ability to visually parse the grammar of written sentences should follow naturally.

This line of reasoning is unsound for several reasons. In the first place, it assumes some sort of abstract identity between speech events and reading as well as a straightforward transfer of speech processing abilities to the reading act. In the second place, even if syntactic competence does transfer unmutilated from normal language to reading, the research on speech production and perception itself is sufficiently ambiguous and incomplete to preclude its usefulness as an exemplar of syntactic processing in reading.

With respect to syntactic processing, we may ultimately discover both similarities and differences between reading and speech processing. The point is that we ought not to be assumptive about the nature of the linguistic overlap between these two areas. With this in mind, the purpose of the following experiment was to provide empirical support for the existence of one particular set of reader strategies as well as to engender some substantive speculation concerning the relationship between reading and speech perception.

A variety of research in both reading and speech perception strongly suggests that the clause is a prominent higher order unit of language processing (Bever, Lackner, and Kirk, 1969; Caplan, 1972; Forster, 1970; Jarvella, 1971; Mehler, Bever, and Carey, 1967; Resnick, 1970; Rode, 1974). If this is true, then it is worth asking how the reader reconnoiters clauses and their boundaries.

Consider the following matched pairs of sentences:

1. (a) John, come here.
   (b) Come here, John.
2. (a) Ben said, "He won't help us dig the pit."
   (b) "He won't help us dig the pit," said Ben.
3. (a) Why can't we eat, John?
   (b) John, why can't we eat?
4. (a) If you can, skip along.
   (b) Skip along, if you can.
5. (a) "Mary can go?" asked Pete.
   (b) "Can Mary go?" asked Pete.
6. (a) "Are you funny!" said Aphrodite.
   (b) "You are funny!" said Aphrodite.

Each of the above pairs are equivalent semantically and in terms of sentence structure complexity, i.e., they mean approximately the same thing, and no presently available method of quantifying sentence structure complexity has the capacity to distinguish between the respective (a)'s and (b)'s. However, this study attempted to substantiate that for some readers such sentence pairs differ in grammatical complexity. In all six cases, (a) is more complex than (b). The complexity in each instance is indicated by the fact that in (a) the underlined punctuation mark is critical to the correct identification of the sentence's grammatical structure. The same is not true of (b), where the punctuation mark is never critical and in some cases is completely redundant. For instance in example 1 (a) the comma cues the reader that "John" is a noun of direct address; but if the comma is deleted or simply not observed, the sentence takes on the appearance of a simple statement with "John" functioning as the sentence subject. This type of linear ambiguity does not exist in 1 (b) if the comma is deleted. A similar analysis can be made for each of the other five pairs.

One of the interesting facets of this phenomenon is that a purely linguistic approach is insufficient for explaining the differences in complexity between the above sentence pairs. In speech, any matched (a) and (b) are equally appropriate and are probably equally easy to produce and perceive, e.g., the transformational history of (a) type sentences is generally no more complex than that of (b) type sentences. However, this equivalence of complexity is not maintained in reading because while (b) type sentences conform to the expectations of readers, (a) type sentences do not. These expectations may be described in terms of canonical-clause strategies.

According to Fodor, Bever, and Garrett (1974), speech perception involves the gathering of information regarding the nature of sentoids, deep structure clauses, which are usually coterminous with surface structure clauses. One heuristic means of gathering such information is the canonical-sentoid strategy, which refers to the hypothesis that the hearer assumes initially that a noun phrase immediately followed by a verb phrase immediately followed by a noun phrase refer, respectively, to the subject, main verb, and object of a common deep structure sentoid.

"Canonical" refers to a fundamental schema which is the clearest or simplest possible, in this case a grammatical schema; and sentoid, within the context of this study, refers to any clause.

There were two canonical shapes or schemas which were included in this study. The first is a statement with the basic grammatical form NP^1-V-(NP^2), where NP^1 refers to the clausal subject, V refers to the main verb of the clause, and (NP^2) refers to an object which is optional, depending upon whether or not the main verb is transitive. The second canonical shape is a question with the basic grammatical form V-NP-X, where V refers to a copula or auxiliary verb, NP refers to the subject of the clause, and X refers to anything that follows the subject of the clause.

Studies investigating grammatical structure in child language and in materials written for elementary school children have revealed that the simple subject-verb-object construction is the most frequently occurring of grammatical patterns (O'Donnell, Griffin, and Norris, 1967; Strickland, 1962). Consequently, the prediction that children search for this pattern is neither profound nor surprising. What is more interesting is the prediction that canonical-clause strategies are so strong in readers that children will seek to confirm their own expectations even in the face of overt visual cues to the contrary, i.e., punctuation. Hence, type (a) clause structures are labeled "non-canonical" and type (b) structures are identified as "canonical"; the latter conform to the reader's grammatical expectations, the former do not.

The experimental hypothesis in this study was that noncanonical clause structures would impede reading because the clause strategies of the subjects would tend to cause them to misprocess the punctuation marks at clause boundaries.

METHOD

Subjects

Subjects were fifty-six third graders from a Title I elementary school in rural Southeastern Ohio. Students participating in special education classes and those identified by their teachers as being unable to read on a first grade level were not included in the sample. Pre-tests of general reading ability, *California Achievement Tests*, Level 2, Form A, 1970 edition, indicated a mean grade level of 3.8 for the fifty-six subjects.

Materials

The stimulus materials in the experiment consisted of two 300 word passages constructed by the investigator and written on a second grade level as estimated by the Fry readability graph. Each passage had two alternate forms. One form of each passage contained noncanonical, type (a), clauses and sentences. The other form contained sentences and clauses which were identical to those of the first form in vocabulary, length, and number of transformations required to derive them, but which were canonical, type (b), structures. The following excerpts are drawn from alternate forms of one of the passages and illustrate the distinction between canonical and noncanonical syntactic structures. Underlined marks of punctuation indicate hypothesized points of syntactic misprocessing in the noncanonical passage.

Noncanonical

> It was a beautiful day. The sun was shining, and the animals were all outside. The zoo had many visitors. It was feeding time when the family went to see the lions. They saw the zookeeper with big baskets of meat. The lions looked hungry.
>
> "Watch them eat, Mother," said Janet.

Canonical

> It was a beautiful day. The animals were all outside, and the sun was shining. The zoo had many visitors. It was feeding time when the family went to see the lions. They saw the zookeeper with big baskets of meat. The lions looked hungry.
>
> "Mother," Janet said, "Watch them eat."

Procedure

The fifty-six subjects were randomly assigned to one of two treatment conditions. Treatment group I read the two passages containing noncanonical, type (a) structures. Treatment group II read the two corresponding passages which contained canonical, type (b), structures. Each group read one passage orally and one silently. The oral reading was timed and also tape recorded so that intonation miscues at the twenty preselected clause boundaries could be evaluated. Criteria used to judge the appropriateness of intonation patterns were based on speech norms defined in Prator and Robinett (1972). Following the reading of each passage, a twelve item multiple choice comprehension test which included both literal and nonliteral questions was administered.

In order to control the overall type I error rate in the experiment, both multivariate and univariate analyses of variance were performed (Finn, 1974). Differences between the two treatment groups were assessed on the following dependent variables: comprehension during silent reading, comprehension during oral reading, speed of oral reading, and intonation miscues at selected clause boundaries.

RESULTS

Subjects reading canonical passages, type (b), scored significantly higher in silent reading comprehension ($M = 9.21$, $SD = 1.45$) than subjects reading noncanonical, type (a), passages ($M = 8.04$, $SD = 1.75$), $F(1, 54) = 7.52$, $p < .01$. In addition, students reading canonical passages generated fewer intonation miscues at clause boundaries ($M = 3.11$, $SD = 2.77$) than students reading the corresponding noncanonical passages ($M = 8.71$, $SD = 2.54$), $F(1, 54) = 62.48$,

$p < .0001$. Nonsignificant differences between treatment groups were obtained on oral reading comprehension and speed during oral reading.

DISCUSSION

The results do not unambiguously support the canonical clause theory since there may be alternative explanations for the canonical/noncanonical phenomenon, e.g., transitional probabilities. However, it would seem that readers, at least by the time they reach third grade, enter into silent reading with definite clause analysis strategies; and these strategies appear to be so strong that they cause children to miss critical marks of punctuation in noncanonical clause structures.

Several conclusions are warranted, the first of which is that grammatical complexity and, hence, readability may depend in part upon clause boundary organizations; and this aspect of grammatical complexity does not reveal itself in assessments of transformational complexity, sentence length, or the frequency with which individual grammatical patterns occur in normal speech. For this reason, clause analysis strategies should be considered as a separate factor in the prediction and explanation of syntactic complexity of written materials.

A second conclusion is that points of punctuation are sometimes minimal language cues which must be extracted from the perceptual input since the alternative is an apparent increase in the probability of syntactic misprocessing. The canonical-clause theory predicts the specific textual conditions under which marks of punctuation are either syntactically critical or redundant.

Superficially, the nonsignificant difference between treatment groups on the variable of oral reading comprehension suggests that syntactic processing relative to clause boundaries is different during oral and silent reading. However, the fact that different passages were employed under the oral and silent conditions could account for the observed discrepancy in treatment effects. It is also possible that the small sample size rendered the experiment insufficiently sensitive to find differences between the treatment groups on oral reading comprehension.

It is not at all clear that clause analysis in reading and clause analysis in speech perception involve identical processes. The listener may indeed possess certain word order expectations, but they could be far stronger in the reader than in the listener because the listener has available to him an alternative system of syntactic cues which increases the redundancy of the perceptual input. These cues are the patterns of pitch, juncture, and stress that combine to form grammatical intonation, the presence or absence of which may be one of the principal distinctions between written and spoken syntax. Consequently, the analysis of grammatical intonation could be a productive area for future psycholinguistic research in reading.

REFERENCES

Baldwin, R. S. *The effects of noncanonical sentences upon the reading comprehension of third grade children.* Unpublished doctoral dissertation, Ohio University, 1976.

Bever, T. G., Lackner, J. R., & Kirk, R. The underlying structures of sentences are the primary units of speech processing. *Perception and Psychophysics,* 1969, *5,* 225-234.

Caplan, D. Clause boundaries and recognition latencies for words in sentences. *Perception and Psychophysics*, 1972, *12*, 73-76.

Finn, J. D. *A general model for multivariate analysis.* New York: Holt, Rinehart and Winston, 1974.

Fodor, J. A., Bever, T. G., & Garrett, M. F. *The psychology of language.* New York: McGraw-Hill, 1974.

Forster, K. I. Visual perception of rapidly presented word sequences of varying complexity. *Perception and Psychophysics*, 1970, *8*, 215-221.

Goodman, K. S. Behind the eye: What happens in reading. In H. Singer and R. B. Ruddell (Eds.), *Theoretical models and processes of reading.* Newark, Delaware: International Reading Association, 1976, 467-491.

Jarvella, R. Syntactic processing of connected speech. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 409-416.

Mehler, J., Bever, T. G., & Carey, P. What we look at when we read. *Perception and Psychophysics*, 1967, *2*, 213-218.

O'Donnell, R. C., Griffin, W. J., & Norris, R. C. *Syntax of kindergarten and elementary school children: A transformational analysis.* Champaign, Illinois: National Council of Teachers of English, 1967.

Prator, C. H., Jr., & Robinett, B. W. *Manual of American English pronunciation.* New York: Holt, Rinehart and Winston, 1972.

Resnick, L. Relations between perceptual and syntactic control in oral reading. *Journal of Educational Psychology*, 1970, *61*, 382-385.

Rode, S. S. Development of phrase and clause boundary reading in children. *Reading Research Quarterly*, 1974, *10*, 124-142.

Strickland, R. G. The language of elementary school children: its relationship to the language of reading textbooks and the quality of reading of selected children. *Bulletin of the School of Education*, Indiana University, 1962, *38*.

CARL T. CLOER, JR.
Furman University, Greenville, South Carolina

SUBVOCALIZATION — ASSET, LIABILITY, OR BOTH?

If educators are to develop a comprehensive theory of reading based on valid research findings, issues related to subvocalization and reading must be resolved. The term subvocalization has been used to describe various degrees of movement or vibration of the vocal chords or facial muscles while reading. Gibson (1975) defines subvocalization as "speech whose range goes from audible sound (e.g., whispering), to movements of speech musculature which must be highly amplified to be detected, to the extreme of speech so implicit that an investigator can make no physical observations" (p. 340).

The issue of subvocalization has elicited the hypothesis by many that subvocalization or implicit speech is not necessary, and therefore is simply a practice that serves as a speed reduction device which prolongs processing during reading. An opposing view is that implicit pronunciation is an aid to meaning and retrieval, and its presence should not be feared as a liability in reading. These opposing views have produced many debates which have led mostly to additional disputations. A need is evident for a review of the literature in an attempt to determine truth for answerable questions.

EARLY DISCUSSION OF SUBVOCALIZATION

Early in the twentieth century, O'Brien (1921) suggested that implicit pronunciation could be prevented by eliminating the oral reading approach to teaching reading. McDade (1937) and Buswell (1945) shared this opinion.

However, Huey (1908) had demonstrated that subvocalization during reading was different from speech. He did a number of experiments with adults who were told to read a novel different ways. They were told to read it "the way you like to read" or to "read aloud," or to "say it all to themselves," or to think of "how it would sound" as they read. Huey discovered that subvocalization was a combination of auditory and motor elements and that reading aloud was 66 per cent slower than reading silently with subvocalization. The inference is that subvocalization was not silent oral reading. Dodge (1907) had already demonstrated that only the beginnings of words were pronounced, and certain words were not pronounced at all during subvocalization. Gibson (1975) suggests that subvocalization differs markedly from oral reading and complete language, even when it is inaudible.

220

TREATMENT TO REDUCE SUBVOCALIZATION

Recent experimental treatment to reduce subvocalization by continuous feedback techniques was reported by Hardyck, Petrinovich, and Ellsworth (1966). The subjects were 17 college students. Surface electrodes were placed near the larynx of the subjects so that muscle movement could be detected. The electrodes produced EMG signals which were converted to audio signals which the readers could hear through earphones. Subjects were instructed to read so as to create no sounds over the phones.

The results demonstrated that one session of the feedback produced complete cessation of vocalization in all subjects. The experiment did not determine what effect the reduction of subvocalization had on the reading comprehension of the subjects.

Hardyck and Petrinovich (1970) conducted a subsequent experiment with 18 freshman students from a remedial English class. The subjects read two essays judged by English instructors to be similar in interest, but varying widely in difficulty of comprehension. EMG recordings were taken from the larynx, the chin and lips, and the right forearm. Subjects were asked to read so as to create no sounds over the earphones. The EMG signals from the larynx, chin and lips areas were most responsive to difficulty level. The EMG activity increased significantly with the difficulty of the task. When the feedback produced a cessation of vocalization, comprehension of the difficult passage suffered. Gibson (1975) suggested that these subjects reduced in comprehension because they had to divide their attention between understanding difficult text and keeping earphones silent. For whatever reason, comprehension suffered.

TASK DIFFICULTY AND SUBVOCALIZATION

One research technique employed to ascertain the relationship between subvocalization and difficulty of the reading task has been to study subvocalization while subjects read text in a language other than their native language. Faaborg-Anderson and Edfeldt (1958) found that when Danish and Swedish adults read a non-native language, EMG's of greater amplitude were recorded. This was especially true if the subjects were unaccustomed to reading foreign prose.

Edfeldt (1960) inserted needle electrodes directly into the speech musculature of students from the University of Stockholm. The subjects read easy and difficult passages, text that was visually clear, and text that was blurred. Clear passages were read with significantly less subvocalizing than blurred text, and electrical activity increased significantly for difficult passages.

Sokolov (1972) had Russian university students translate English texts of varying difficulty into Russian. He conducted EMG studies on the translation task with electrodes placed on the tongue and lower lip muscles. The more difficult translations resulted in more electrical activity in the speech musculature. Furthermore, when the subjects were told to reread the texts more attentively, there was a significant intensification of motor speech excitation.

These studies consistently show that when the reading tasks become more difficult by the nature of the prose, the rate of subvocalization increases.

However, most of the research has been conducted with adult subjects. Very little is known about the relationship between the difficulty of the reading task and subvocalization among children.

DEVELOPMENTAL COURSE OF SUBVOCALIZATION

Pomerantz (1971) suggested that subvocalization might be viewed as a natural developmental reinforcement mechanism. Anderson and Dearborn (1952) suggested that any reader should feel free to call on the reservoir of deeper meaning of the spoken language when encountering difficulty in understanding reading material. One implication is that comprehension may be mediated by speech to a greater degree among immature readers than among more sophisticated and facile readers.

McGuigan and Bailey (1969) retested subvocalization among elementary-school age children after two and three year periods. The original whispering had completely disappeared over the two- to three-year period. After the second year, EMGs from the lips and chin had reduced significantly.

Cleland, Laffey, and Anderson (1969) found a negative correlation between subvocalization and scores on various language skill tests. They concluded that a natural decrease in subvocalization accompanied growth toward maturity in language skills. They hypothesized that as the reader meets with less frustration and is more assured, the resultant subvocalization should slowly decrease.

There is a veritable dearth of data concerning the course of subvocalization as a child becomes a more skillful reader. Much longitudinal research is needed that would provide data divided according to chronological age, grade level, reading ability, and difficulty of the text.

THE EFFECTIVENESS OF SUBVOCALIZATION EXTINCTION TRAINING

Educators involved with college and adult reading are familiar with the term "phrase reading". Advocates of phrase reading stress the importance of reading groups of words or 'complete thoughts' rather than one word at a time. Much evidence has been presented to the effect that mature readers extract information in the largest units of which they are capable. Gibson (1975) stated that phrase reading is an achievement which one could hardly fail to find praiseworthy.

College and adult reading improvement courses often involve the teaching of phrase reading with one-third and one-fourth second fixations on each phrase. The implication is that readers cannot possibly be "saying words in their heads" while reading three words in one-fourth second fixations. But the question is whether or not such training is effective and lasting.

Voluse (1973) compared the effects of speed reading instruction with and without subvocalization extinction training. One group from a sample of 29 college subjects was assigned to a treatment providing training in rapid discrimination of phrases. The other group received this training plus audio-feedback from the thyroid area. Voluse reported "substantial" reductions in thyroid subvocalization and increases in speed for each group. No statistical tests, however, were reported.

McGuigan (1970) found that training to reduce subvocalization was ephemeral. After the tones were removed, subvocalization returned to the pretest levels. There was no evidence that any of the subjects reduced subvocalization without the external reinforcement. When subvocalization was lowered during training, reading rate increased and there appeared to be no loss in comprehension. Edfelt (1960) had demonstrated an inverse correlation between reading rate and amount of subvocalization.

Aarons (1971), however, found that small groups of adults previously determined to be high or low subvocalizers were all responsive to feedback training, and the reductions in subvocalization lasted beyond the training trials. Hardyck et al, (1966) also found no evidence of subvocalization after one and three month periods. These studies reveal a discrepancy as to the ephemeral effects of subvocalization extinction training. The discrepancy in these studies as to the ephemeral effects may be explained by the fact that McGuigan (1970) used subjects 7 to 19 years of age with no analysis by age. Hardyck et al, (1966) and Aarons (1971) used college and adult readers.

SUMMARY

What conclusions or recommendations can be drawn from the literature reviewed? It appears that the degree of subvocalization that occurs while reading depends on the reader's text and the reader's skill, as well as specific training strategies. As the text becomes difficult to understand, subvocalization increases. As readers become more sophisticated, it decreases. Auditory feedback of speech muscle activity has a rapid and persistent effect in reducing subvocalization.

There seems to be an inverse correlation between reading rate and amount of subvocalization. However, there is little or no proof that subvocalization causes poor reading ability or that the absence of subvocalization is a sign of efficient reading. Subvocalization as a reading disorder or as a symptom of a disorder is not well substantiated. Subvocalization has been recorded in many situations where it is inappropriate to refer to a reading "disorder". Remediation may be appropriate in instances where conscious subvocalization is seriously limiting the rate of reading simple materials.

There is a veritable need for data concerning the course of subvocalization as the child becomes a skilled reader. The inconsistent evidence to data suggests that implicit speech may be a natural developmental reinforcement mechanism, a mechanical device that slows the perusal of text, or a comprehension aid by adding the redundancy of another modality.

If implicit speech becomes unnecessary for reading any type of material at a certain stage in the development of a reader, there are important implications for improving reading rate. If there are certain pedagogical techniques in teaching reading that tend to produce more subvocalization over longer periods of time, these may need more careful scrutiny in the formative years of reading instruction.

But to answer the original question posed as to whether subvocalization is an asset, liability, or both, it appears that the advantages of implicit speech may outweigh the disadvantages. Subvocalization is a liability only in terms of reading speed. Its relation to comprehension is not well established but there is no evidence that it interferes with comprehension. Flexibility in reading rate is a more desirable goal than a fixed fast rate, and comprehension is more desirable than rapid speed. Until additional data has been analyzed, it appears that subvocalization should be viewed as something other than a definite liability, an asset perhaps, or simply something that happens quite naturally for certain readers in certain stages or circumstances.

REFERENCES

Aarons, L. Subvocalization: Aural and EMG feedback in reading. *Perceptual and Motor Skills,* 1971, *33,* 271-306.

Anderson, I. H. & Dearborn, W. F. *The psychology of teaching reading.* New York: Ronald Press, 1952.

Buswell, G. T. Perceptual research and methods of learning. *The Scientific Monthly,* 1947, *64,* 520-521.

Cleland, D. L., Laffey, J., & Anderson, J. W. *Vocalism in silent reading* (Report of Project No. BR-5-0374). Washington, D. C.: United States Department of Health, Education, and Welfare, 1968.

Dodge, R. An experimental study of visual fixation. *Psychological Review Monograph Supplement,* 1907, *8,* No. 4.

Edfeldt, A. W. *Silent speech and silent reading.* Chicago: University of Chicago Press, 1960.

Faaborg-Anderson, K. & Edfeldt, A. W. Electromyography of intrinsic and extrinsic laryngeal muscles during silent speech: Correlation with reading activity. *Acta Otolaryngologia,* 1958, *49,* 478-482.

Gibson, Eleanor J., & Levin, H. *The psychology of reading.* Cambridge, Mass.: The MIT Press, 1975.

Hardyck, C. D., & Petrinovich, L. F. Subvocal speech and comprehension level as a function of the difficulty level of reading material. *Journal of Verbal Learning and Verbal Behavior,* 1970, *9,* 647-652.

Hardyck, C. D., Petrinovich, L. F., & Ellsworth, D. W. Feedback of speech muscle activity during silent reading: Rapid extinction. *Science,* 1966, *154,* 1467-1468.

Huey, E. B. *The psychology and pedagogy of reading.* New York: Macmillan, 1908. Republished, Cambridge, Mass.: MIT Press, 1968.

Klapp, S. T., Anderson, W. G., & Berrian, R. W. Implicit speech in reading, reconsidered. *Journal of Experimental Psychology,* 1973, *100,* 364-374.

McDade, J. E. A hypothesis for nonoral reading: Argument, experiment, and results. *Journal of Educational Research,* 1937, *30,* 489-503.

McGuigan, F. J. Covert oral behavior during the silent performance of language tasks. *Psychological Bulletin,* 1970, *74,* 309, 326.

McGuigan, F. J. & Bailey, S. C. Longitudinal study of covert oral behavior during silent reading. *Perceptual and Motor Skills,* 1969, *28,* 170.

O'Brien, J. A. *Silent reading.* New York: Macmillan Co., 1921.

Pomerantz, H. Subvocalization and reading. *The Reading Teacher,* 1971, *24,* 665-667.

Sokolov, A. N. *Inner speech and thought.* New York: Plenum Press, 1972.

Voluse, C. R., III. *Depressing adult subvocalization behaviors utilizing direct-audio-feedback and/or reading efficiency training.* Unpublished doctoral dissertation, University of Virginia.

BONNIE J. F. MEYER
Arizona State University, Tempe, AZ

ORGANIZATION IN PROSE AND MEMORY: RESEARCH WITH APPLICATION TO READING COMPREHENSION

In the last five years there has been a convergence of interest from various disciplines focusing on the use of organization in prose by a reader. Educators concerned with teaching reading have particular interest in this area as it relates to the practical problem of reading comprehension. In addition, there has been a recent surge in research and theoretical effort in this area by cognitive science, a domain including psycholinguistics, cognitive psychology, and educational psychology as well as artificial intelligence. This work in cognitive science has been influenced substantially by recent advances in linguistics, rhetorics, and philosophical logic.

Educators have long recognized that reading comprehension involved skill in following the structure of a passage and recognizing the writer's purpose (Carroll, 1972; Davis, 1941). Recent books and manuals encourage teachers to teach students to look for the writer's organization or pattern in order to increase their efficiency in identifying the message of a passage and to increase their retention of it (Dechant, 1970; Herber, 1970; McGuire & Bumpus, 1971; Robinson, 1975). Recommendations stress the importance of teaching students to relate ideas together within and between sentences (McElwee, 1974; Strang, McCullough, & Traxler, 1967).

Recommendations for teachers to identify and use the organization in prose tend to be based on common sense notions that may or may not prove valid. What is needed is a theoretically based procedure for identifying the structure of a passage and an adequate theory of learning and memory from discourse to explain the function in learning of this structure in prose. Although a complete version of such a theory will not be immediately forthcoming, research in cognitive science aimed at this task has made substantial progress in the last few years; educators should keep in close contact with the developments in this area.

Until recently, educators could count on little assistance from psychology or linguistics specifying processes and procedures involved in higher order comprehension skills. The linguistics of Bloomfield and behavioristic theory of language behavior did not deal with meaning nor the function of information in prose. The work of Chomsky (1957) dealing with meaning through the concept of deep structure revolutionized linquistics and spurred the development of psycholinguistics and cognitive psychology.

Fillmore (1968) went further in his probe of meaning. He rejected the notion of the subject-predicate division as the framework for the underlying structure of a sentence. In the place of Chomsky's syntactic deep structure, Fillmore postulated a semantic deep structure based on the notions of case relationships or the role of noun phrases in sentences. The deep structure of a proposition (simple sentence) consists of a predicate (usually the verb) and one or more arguments (noun phrases); the predicate and its arguments are related together with case relationships which specify the function of each argument in the action or state.

Other linguists (Chafe, 1970; Grimes, 1975; Halliday, 1967) as well as a revision by Fillmore (1971) list case-like semantic descriptions that vary in number and label of cases from Fillmore's original cases. As Grimes (1975) points out, there is consensus among the theorists on the general framework of case grammar and differences reflect different attempts to specify the details.

The influence of case grammar in cognitive science has been substantial. Models of semantic memory are based on case grammar (Fredericksen, 1975; Kintsch, 1974; Meyer, 1975; Norman & Rumelhart, 1975; Schank, 1972). In addition, a number of investigations have pointed to the psychological reality of case grammar (Isakson, 1975; Kintsch, 1972; Shafto, 1973; Suci & Hamacher, 1972; Wilson, 1975; Yekovich & Walker, 1976). Research examining the processing of different case relations in sentences continues in a search to understand the effects of case on comprehension and memory.

A current theory re-emerging in cognitive science from the seminal works of Kant (1787) and Bartlett (1932) is that of schema theory (Anderson, in press; Rumelhart & Ortony, in press; van Dijk & Kintsch, in press). Similar theories from workers in artificial intelligence (Minsky, 1975; Winograd, 1975) have used the term frame for schema and Schank (Schank & Abelson, 1975) has used the term plan. Rumelhart and Ortony (in press) present a detailed account of schema theory. Briefly, schemata are cognitive templates or patterns. They have variables which can be filled with specific content, and they can be embedded upon each other. There are schemata for individual predicates and types of prose.

A recent paper by van Dijk and Kintsch (in press) presents an explanation for the interrelationships between the organization of narratives and schema theory. Discourse is divided into two levels of meaning. The first is the micro-level; it corresponds to case grammar analysis of sentences with added connectives and quantifiers to specific intersentence and other intrasentential relations. The second level of meaning is that of parts of discourse or of the discourse as a whole and is labeled the macro-structure. The macro-structure of a passage can be either explicit or implicit. A summary of a passage is similar to what van Dijk envisions as macro-structure. Macro-structures relate to micro-structures through rules of deletion and subordination. The nonessential information and repetitive information of micro-structures are not included in the macro-structure. In addition, macro-structures include generalizations and superordinate ideas which subsume information found in the micro-structure; these generalizations and superordinates are not in the micro-structure.

According to the position posited by van Dijk and Kintsch, a reader approaches a narrative with a narrative schema in mind. (Schemata of stories have been specified by Rumelhart (1975) expanding on the earlier work of Propp (1958). For example, the highest level of a story schema is setting and episode; episode can be further broken down into variables.) The narrative schema that a reader brings to a narrative has its variables filled in with appropriate information from the macro-structure inferred or explicit in the text.

Rumelhart (in press) clearly and simply states a model of comprehension related to van Dijk and Kintsch's position. He explains that the process of comprehension is the same as the process of choosing and verifying conceptual schemata to account for the text to be understood. A schema accounts for a passage whenever the passage can be taken as an instance of the general concepts represented by the schema.

Schema theory views recall as a reconstructive process. In processing the information from a passage for memory, an adequate schema or schemata from a person's memory is identified and employed. The information of the macro-structure that fits the schema's variables is stored in memory along with details low in the structure (micro-level) which catch the reader's attention (Meyer, 1975). In recall the reader locates in memory the schemata used in processing the passage and the stored traces from the passage; these traces are a subset of the information from the passage with the prevalence of traces corresponding to information high in the structure at the macro-structure level. Recall involves utilizing the schemata to reconstruct the passage with the information stored in memory from the passage and information added to fit the schemata. Thus, recall protocols from passages with structures that closely fit the variables of existing schemata consist of a subset of the ideas in the original passage with few intrusions (Meyer, 1975). In contrast, recall protocols from passages with obscure macro-structure fitting no well-defined schemata in the culture contain many intrusions and distortions (Bartlett, 1932). Anderson et al. (1976) explain that distortions and intrusions only occur when there is a lack of correspondence between the schemata of the passage and the schemata employed by the reader to assimilate the passage.

The findings from the empirical investigations with prose of Meyer (1975, in press) can be interpreted through schema theory and the macro-structure and micro-structure model. Meyer's analysis of prose is based on Fillmore's (1968) case grammar and Grimes' (1975) semantic grammar of propositions. In addition to the lexical predicates (verbs) and their arguments of case grammar found in Meyer's analyses, there are rhetorical predicates. These rhetorical predicates consist of a finite number of labels which classify and describe the relationships, particularly intersentential and inter-paragraph relations, found in prose. An analysis of a passage with the semantic grammar of propositions yields a hierarchically arranged content structure that diagrams all the ideas in a passage and contains labels which specify the relationships among the ideas.

Research on learning and memory with this content structure shows that height of information in this content structure is related to its acquisition, recall and retention. Ideas located high in the structure are better remembered after reading or listening to a passage than ideas low in the structure (Meyer & McConkie, 1973; Meyer, 1975; Meyer, in press). These findings hold for different age (sixth grade, undergraduate and graduate students) and ability levels (above, at, and below grade level in reading) and recall tasks.

Information high in Meyer's content structure corresponds to van Dijk's macro-structure, while information lower in the structure corresponds with his micro-structure. Rhetorical predicates at the top levels of the content structure need not be explicit. One study Meyer (1975) compared the recall of passages with their rhetorical structure made explicit versus those where these relationships were not explicitly stated. The results indicated that explicit statement of the rhetorical structure facilitated recall. Other variables were considered at the same time confounding the results such as explicit statement by a writer of pointer words stating "this is an important point"; further research is necessary in this area.

Another study by Meyer (1975) showed that passages with the same hierarchical structure, case relations, and rhetorical predicates, but different content produced patterns of recall that were similar for the passages ($r = .55$) and particularly for the top half of the content structure ($r = .83$). This correlation was not due simply to the height in the structure; when relational similarities of case and rhetorical relations were removed and only the height factor remained, the correlation for the total passage dropped from .55 to .22. A plausible explanation for the similar patterns of recall from the passages with the same structure, but different content is that the same schema was provided by subjects to process both passages. The top rhetorical predicate for both passages was a response rhetorical predicate. A response rhetorical predicate relates a problem (or question) to a solution (or answer). Thus, the macro-structure or top-level structure of both passages could be accounted for by a response schema. A response schema is commonly used by writers of scientific articles.

In a recent study, Meyer and Freedle (in press) studied the effects of different top-level rhetorical structures on what people remember from passages. Four top-level rhetorical structures were examined: *response* rhetorical predicate, *adversative* rhetorical predicate which relates what did happen to what did not happen, *covariance* rhetorical predicate which relates an antecedent to a consequent, and *attribution* rhetorical predicate which relates a collection of attributes to an event or idea. The same information was presented in four passages differing only in their top-level rhetorical structures. The study clearly showed that some top-level rhetorical structures (adversative and covariance) are superior to others (attribution) in facilitating recall and retention of identical information.

Identification of the top-level structure of schema employed by the writer to organize his content can be of great assistance to the reader in processing information from text. It appears that information to be learned and retained by a reader must be organized into a system of logical relationships by the reader. In reading a passage, a reader can decide either to use the same type of schema employed by the author or provide another different type of schema for organizing the information for processing and storage in memory.

In most school situations and many learning situations outside of school, a person wants to know exactly what an author said. The reader needs to pick up as much of the information presented as possible as well as retaining the author's message or main ideas. For reading situations of this type, it seems reasonable that the most efficient strategy for a reader is to utilize the organization of the writer and store the information in the passage for memory in the same type of schema as that used by the writer. This strategy is posited to be more efficient because the reader does not have to search his memory for an alternative and appropriate schema nor does he have to reorganize the ideas in the passage to fit this different schema while reading; instead, he saves processing time by utilizing the same schema as that of the author and organizing the information in the same way as the author. Some data gathered (Meyer & Freedle, 1976) indicate that when readers or listeners utilize the organization of the author, they can remember more of the ideas of the passage than readers who have not used the author's organization; in this study the exposure time for both groups was held constant.

Of course, there are times when using schema of the same type as the author is not efficient. One such time would be when the reader's purpose is simply scanning an article for a specific detail. Another time would be when the reader disagrees with the schema of the author. Data relating to this alternative strategy have been collected (Meyer & Freedle, 1976) from school teachers reading a passage with a problem-solution schema. The solution in this passage was "immediate dismissal of athletic coaches by school boards"; the teachers tended not to report this solution in their recall protocols and did not organize their recall in terms of the author's problem-solution format, but instead organized their recall protocols with adversative or attribution schemata. In addition, on occasions a reader must integrate information by a number of authors on one topic. The first reading of the articles would probably be most efficient with the proposed strategy of utilizing the same type of schema as that of the author, but later integration and comparison would require restructuring with different schemata provided by the reader.

Thus, although there are certain situations when using the same schema as that used by the writer is not desirable, it appears to be the most efficient strategy for most school learning situations. It would seem to be important for students to be able to identify the top-level rhetorical predicates or writer's schema found in prose. A study in which the author is currently involved examines whether or not junior high school students who differ in their reading comprehension skills differ in their use of the same type of schema as that provided by the author. If students with poor comprehension do not identify and utilize the organization in the text, training programs should be designed to develop such skills in readers. A second study is underway to provide data for determining the type of training most effective in such a program. This on-going study deals with whether assistance in identifying the top-level organization of a passage facilitates recall of the passage. A number of different ways to facilitate identification of a writer's schema are examined.

Research dealing with the function of organization in prose and memory has practical implications for educators. A pedagogical application would be explicit teaching of how to identify the schemata used by an author and commonly used schemata, such as response, covariance, adversative and story structure. Another application would be for teaching writing skills. Writers must provide a well-organized text with clearly identified schemata and realize that some schemata are more effective than others in processing information. Meyer (in press) can be referred to for further discussion of the uses of content structure in educational practice as a diagnostic tool, guide for writing main idea and detail questions, and a pedagogical tool.

REFERENCES

Anderson, R. C. The notion of schemata and the educational enterprise. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Lawrence Erlbaum Associates, in press.

Anderson, R. C., Reynolds, R. E., Schallert, D. L., & Goetz, E. T. *Frameworks for comprehending discourse* (Tech. Rep. No. 12). Urbana, Ill.: University of Illinois, Laboratory for Cognitive Studies in Education, 1976.

Carroll, J. B. Defining language comprehension: Some speculations. In J. B. Carroll & R. O. Freedle (Eds.), *Language comprehension and the acquisition of knowledge.* Washington, D.C.: V. H. Winston, 1972.

Chafe, W. L. *Meaning and the structure of language.* Chicago, Ill.: University of Chicago Press, 1970.

Chomsky, N. *Syntactic structures.* The Hague: Mouton, 1957.

Davis, F. B. *Fundamental factors of comprehension in reading.* Unpublished doctoral dissertation, Harvard University, 1941.

Dechant, E. *Improving the teaching of reading.* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.

Fillmore, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory.* New York: Holt, Rinehart and Winston, 1968.

Fillmore, C. J. Some problems for case grammar. In R. J. O'Brien (Ed.), *Report of the twenty-second annual round table meeting on linguistics and language study.* Washington, D.C.: Georgetown University Press, 1971.

Fredericksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology,* 1975, *7,* 371-458.

Grimes, J. E. *The thread of discourse.* The Hague, Holland: Mouton and Co., 1975.

Halliday, M. A. K. Notes on transitivity and theme. *Journal of Linguistics,* 1967, *3,* 37-81, 199-244.

Herber, H. *Teaching reading in the content areas.* Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1970.

Isakson, R. *Case role identification and sentence meaning processing.* Unpublished doctoral dissertation, Cornell University, 1975.

Kintsch, W. Notes on the structure of meaning in memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press, 1972.

McElwee, M. Techniques for improving comprehension in science. *Reading in the middle school.* Newark, Dela.: International Reading Association, 1974.

McGuire, M., & Bumpus, M. *The Croft inservice program: Reading comprehension skills.* New London, Conn.: Croft Educational Services, 1971.

Meyer, B. J. F. *The organization of prose and its effects on memory.* Amsterdam: North-Holland Publishing Co., 1975.

Meyer, B. J. F. The structure of prose: Effects on learning and memory and implications for educational practice. In R. C. Anderson, R. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale: Lawrence Erlbaum Associates, in press.

Meyer, B. J. F., & Freedle, R. *The effects of different discourse types on recall.* Princeton, N.J.: Educational Testing Service, in press.

Meyer, B. J. F., & McConkie, G. W. What is recalled after hearing a passage? *Journal of Educational Psychology,* 1973, *65,* 109-117.

Minsky, M. A framework for representing knowledge. In P. Winston (Ed.), *The psychology of computer vision.* New York: McGraw Hill, 1975.

Norman, D. A., Rumelhart, D. E., & the LNR Research Group. *Explorations in cognition.* San Francisco: W. H. Freeman and Co., 1975.

Propp, V. *Morphology of the folktale* (L. Scott, trans.). Bloomington, Ind.: Research Center in Anthropology, Folklore, and Linguistics, 1958.

Robinson, A. *Teaching reading and study strategies: The content area.* Boston: Allyn and Bacon, Inc., 1975.

Rumelhart, D. E. Notes on a schema for stories. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding.* New York: Academic Press, 1975.

Rumelhart, D. E. Understanding and summarizing brief stories. In D. LaBerge & J. Samuels (Eds.), *Basic processes in reading: Perception and comprehension.* Hillsdale, N.J.: Lawrence Erlbaum Associates, in press.

Rumelhart, D., & Ortony, A. The representation of knowledge in memory. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge.* Hillsdale, N.J.: Lawrence Erlbaum Associates, in press.

Schank, R. C. Conceptual dependency: A theory of natural language understanding. *Cognitive Psychology,* 1972, *3,* 552-631.

Schank, R. C., & Abelson, R. P. Scripts, plans and knowledge. *Advance Papers of the Fourth International Joint Conference on Artificial Intelligence.* Tbilisi, Georgia, USSR, 1975, 151-157.

Shafto, M. The space for case. *Journal of Verbal Learning and Verbal Behavior*, 1973, *12*, 551-562.

Strang, R., McCullough, C., & Traxler, A. *The improvement of reading*. New York: McGraw-Hill Book Company, Inc., 1967.

Suci, G. J., & Hamacher, J. H. Psychological dimensions of case in sentence processing: Action role and animateness. *International Journal of Psycholinguistics*, 1972, *1*, 34-48.

van Dijk, T., & Kintsch, W. Cognitive psychology and discourse: Recalling and summarizing stories. In W. U. Dressler (Ed.), *Trends in text-linguistics*. New York: de Gruyter, in press.

Wilson, S. J. *Some influences of semantic structure on the recall of sentences*. Unpublished doctoral dissertation, Cornell University, 1975.

Winograd, T. Frame representations and the declarative-procedural controversy. In D. G. Bobrow & A. M. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. New York: Academic Press, 1975.

Yekovich, F. R., & Walker, C. *On the formation and comprehension of linguistic argument repetitions.* Arizona State University, paper in progress.

DOROTHY M. HANSEN
Brigham Young University

A DISCOURSE STRUCTURE ANALYSIS OF THE COMPREHENSION OF RAPID READERS

This paper reports the results of two studies designed to measure qualitative differences in comprehension between readers trained in the Wood method and normal readers.

Spache (1962), Ehrlich (1963), Tinker (1962), Carver (1971, and Graf (1973) tested Wood readers against controls by having both groups read the same material once, timing the rate, and measuring comprehension with a multiple-choice test. All studies reported that the mean comprehension of speed readers decreased as rate increased, but no researcher classified his test items as to type of comprehension required for answering correctly. Liddle (1965) obtained mixed results by varying the method of reading to allow for preview and postview, and using items calling for critical reflection, inference, or fact recognition.

In a more sophisticated approach to comprehension, language theorists Carroll (1972), Grimes (1972), Pitkin (1973), Meyer (1974), and Packard (1975) demonstrated the interconnectedness of structure and comprehension, and provided the methodology for measuring what parts of discourse are recalled.

The studies reported here were designed to vary reading conditions and method of analysis to determine qualitative differences between the comprehension of Wood readers and normal readers (i.e., were rapid readers getting only the main ideas?).

METHOD

In the first study, freshmen honor students were randomly assigned to one of two experimental groups or to one control group. All groups were taught a basic course in composition and rhetoric with heavy emphasis on the Christensen generative rhetoric. In addition, the experimental groups received one hour of instruction per week for seven weeks in the Wood method of rapid reading.

Comprehension was measured by comparing pre- and posttest written recalls. On the pretest students were given 13 minutes (325 words per minute) to read the first one-third of a chapter on thinking from a psychology textbook. For the posttest the students read the last two-thirds of the same chapter, again in 13 minutes (772 words per minute) for a rate-power test. After each reading students were instructed to write down all they could remember, and were given all the time they needed. Students were not allowed to reread if they finished early. All the experimental students finished the material before the time was up.

The second study was conducted the following semester and was the same as the first in all respects but four: (1) the students read approximately half the chapter for the pretest and the other half for the posttest; (2) the reading time was extended from 13 to 19 minutes, requiring 405 words per minute to finish the pretest, and 525 words per minute to finish the posttest; (3) students were instructed to reread if they finished early, using all their time in an efficiency-power test; (4) after the students had written all they could, they were given an objective test designed to elicit concept application, paraphrase, and verbatim recall comprehension (Anderson, 1972). All the experimental students read the material at least twice.

For the discourse analysis study, 20 paired recalls were randomly selected from each semester, 10 from the experimental and 10 from the control group, for a total of 40 students. Incomplete date required the control group second semester to be reduced to 7, for a total of 37 students with pre- and post paired recalls.

ANALYSIS

The chapter from a psychology textbook which was used for this study was analyzed according to Packard's (1975) notational adaptation of Pitkin's (1973) theory.[1] Packard established six discourse relationships, or operations; coordination, temporal, comparison/contrast, reason, emphasis, and elaboration. He defined these relationships by using extensive examples for selecting which relationship existed between units of discourse. For the sake of uniformity, Packard used clauses as his basic unit of discourse, and defined a hierarchical tree structure to represent the information units and their relationships to each other. An example appears in Figure 1.

Using the extensive examples provided, five raters spent a total of two weeks analyzing different parts of the chapter. The raters were given practice in analyzing parts of discourse according to the connectives listed by Packard. When they felt sufficiently practiced, part of the chapter used in the study was analyzed by all five raters. A correlation of .93 was obtained on this trial by using the correlation ratio eta, where

$$\text{eta} = \frac{\text{ss between groups}}{\text{ss total}}$$

Periodically during the two-week analysis period, interrater reliability checks were run, to be sure the reliability figure remained high. All intermediate checks yielded correlations above .90, satisfying the demands of the study. (This analysis, covering 104 pages, is available from the author.)

The recalls, identified only by number, were judged against the discourse analysis of the chapter to determine which information units were recalled and which relationships governed the information units recalled. Recall was credited if a clause was accurately recalled, or if a clause was paraphrased with the essen-

---

[1] (See Hansen, 1975, for a detailed analysis of the Packard method and rationale for using it instead of others.)

Figure 1.

1 The revival of interest in rhetoric in the past twenty years, and in particular the "new rhetorics" that have emerged during the 1960's, had led teachers of written composition to hope that they would soon have a workable rhetoric of the whole essay.

2a This rhetoric, we hoped, would explain the essentials of how complete essays are made to achieve their purposes,

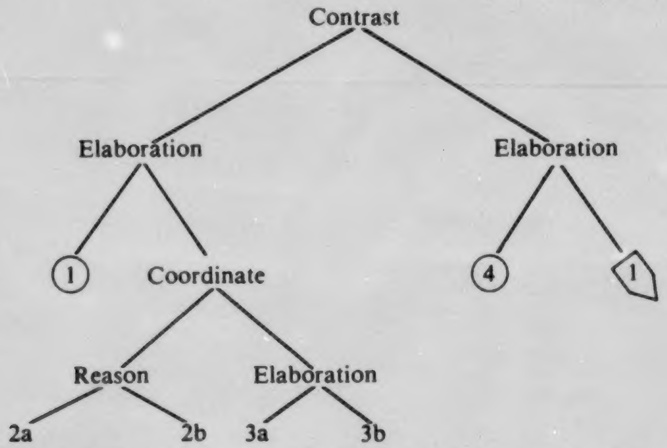
2b so that writers could learn reliably how to choose and assemble materials in order to produce a successful essay.

3a We hoped, too, that this rhetoric of the whole essay would be "generative";

3b that is, we hoped that it would help writers learn how to discover the materials they need in order to say what they want, and would help them identify ways to get their pieces started and carry them on to their conclusions.

4 But thus far articles and proposals in the professional literature on composition have not offered us the generative rhetoric we are seeking.

(Richard Larson, 1971)



An illustration of the Packard diagram of discourse relationships. The numbers on the tree correspond to the numbers in the prose passage, with the kinds of relationships joining the discourse indicated above each group.

tial meaning accurately preserved. A relationship was credited (1) whenever both information units immediately under an operation were recalled; (2) when one unit at the first level below the operation and another unit at the second level below the operation, on another branch, were recalled; and (3) when two units each at the second level below the relationship to be credited were recalled. A cluster was defined as consisting of two or more information units plus the relationships governing those units.

The raters wrote down which information units in which paragraphs were recalled, then transferred those lists to the structural analysis of the chapter for summary and analysis.

RESULTS

An analysis of variance and covariance used to measure the differences in experimental and control pre- and posttest scores from the first semester study showed no significant differences between groups on total information units recalled, total operation units governing recall, number of information units recalled which were governed by elaboration or comparison/contrast types of operations, total number of clusters, or the total number of main ideas recalled.

The control group showed significantly greater gains in the number of units recalled which were governed by coordination (.05).

The results from the second semester study were quite different from those of the first semester. These students had been given the same amount of material on pre- and posttests, and the same amount of time. The rapid readers were thus able to read the material more than once. The experimental groups showed significantly greater gains in total units recalled (.01); total operations governing recall (.05), with greater gains on information units controlled by reason (.05), comparison/contrast (.001), and coordination (.05) relationships; the total number of units in clusters (.05); and the number of clusters (.001). There were no significant differences on recall of information units governed by elaboration, temporal, or emphasis operations; average cluster size; or number of main ideas recalled.

Surprisingly, neither group showed any significant differences in comprehension on the objective test in any category, concept application, paraphrase, or verbatim recall. A discourse analysis of the test failed to account for it's inability to discriminate the comprehension differences between the two groups.

DISCUSSION AND SUMMARY

These studies provided insight into several areas of controversy existing between opponents and proponents of rapid reading.

First, rapid readers are not just skimming for main ideas. In neither study was there a significant difference between rapid and normal readers in the recall of main ideas. Rather, the comprehension was quite evenly distributed for both groups among main ideas and supporting detail.

Second, it appears that students get better comprehension from reading a passage twice at high rates than once at slow rates. Given the same amount of time to master a passage, rapid readers significantly outperformed normal readers. Rapid readers recalled more information units, and more of these units were recalled in relationship to other units The size of the clusters did not differ from experimental to control, however, remaining around six units. This result probably provides more evidence for Miller's 7 ± 2 theory of memory (Miller, 1956), in which Miller proposed that the capacity of short-term memory is seven items (or "chunks" of information), plus or minus two.

Third, and perhaps most important, the method of reading produced qualitative differences in comprehension not explainable simply as quantitative differences. For example, an inspection of the raw data revealed that 70 percent of the rapid readers recalled paragraph 63, while this same paragraph was not mentioned by any slower reader, although slow readers recalled information from paragraphs on both sides of 63. Paragraph 63 contains three examples of deductive logic in formal syllogistic form, with the examples indented and the parts of the syllogism clearly identified. The obvious difference in the paragraph is its surface appearance. It is possible that rapid readers respond more quickly to these visual clues than slower readers. This visual clue might also account for the unexpected gain on reason nodes demonstrated by the rapid readers.

A second indication of qualitative differences is an apparent selective change across the relationships governing information recalled for both experimental and control groups. For instance, under both conditions, slower readers recalled more information units governed by temporal relationships. Also, the constellation of increments was different under both conditions for rapid and slow readers. In the first study slow readers recalled more information governed by coordination, temporal, and emphasis relationships. In the second study, rapid readers recalled more information governed by reason, comparison/contrast, and coordination relationships, those relationships which might logically be assumed to indicate "gestalt" comprehension.

If, as these studies suggest, the method of reading determines what is recalled, then this ought to be a well-defined variable in all memory and comprehension studies dealing with natural language. As yet, it does not appear to have been considered at all. Only as researchers move toward delineating all the complex factors involved in reading — structural, emotional, methodological, and individual — will be a satisfactory theory of comprehension emerge.

REFERENCES

Anderson, R. C. How to construct achievement tests to assess comprehension. *Review of Education Research,* 1972, *42,* 145-170.

Carver, R. P. *Sense and nonsense in speed reading.* Silver Spring, Maryland: Revrec Publishing Company, 1971.

Christensen, F. A generative rhetoric of the sentence. *College Composition and Communication,* 1963, *14,* 155-161.

Ehrlich, E. Opinions differ on speed reading. *NEA Journal,* April 1963, 43-44.

Freedle, R. O. & Carroll, J. B. (Eds.). *Language comprehension and the acquisition of knowledge.* New York: Wiley, 1972.

Graf, R. G. Speed reading: Remember the tortoise. *Psychology Today,* December 1973.

Grimes, J. E. *The thread of discourse.* Ithaca: Cornell University, 1972.

Hansen, D. A discourse structure analysis of the comprehension of rapid readers (Doctoral dissertation, Brigham Young University, 1975). *Dissertation Abstracts International,* 1976, *36,* 7293-A. (University Microfilms No. 76-9839)

Larson, R. Toward a linear rhetoric of the essay. *College Composition and Communication,* 1971, *22,* 140-146.

Liddle, W. An investigation of the Wood reading dynamics method (Doctoral dissertation, University of Delaware, 1966). *Dissertation Abstracts International,* 1966, *27,* 605-A. (University Microfilms No. 66-5559)

Meyer, B. J. F. *Idea units recalled from prose in relation to their position in the logical structure, importance, stability and order in the passage.* Unpublished master's thesis, Cornell University, 1971.

Miller, G. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review,* 1956, *43,* 81-97

The Organization of Prose and Its Effect on Recall (Doctoral dissertation, Cornell University, 1974). *Dissertation Abstracts International,* 1974, *35,* 3529-A. (University Microfilms No. 74-26291)

Packard, D. J. *Discourse — A practical introduction for readers and writers.* Unpublished class manual, Brigham Young University, 1975.

Pitkin, W. L., Jr. A pedagogical model of discourse structure (Doctoral dissertation, University of Southern California, 1973). *Dissertation Abstracts International*, 1973, *34*, 4279-A. (University Microfilms No. 73-31378)

Spache, G. D. Is this a breakthrough in reading? *The Reading Teacher*, January 1962, 258-263.

Tinker, M. A. *Uses and limitations of speed reading programs in schools.* Report given at Conference on Speed Reading, University of Delaware, March 1962.

JAMES T. FLEMING
FRED W. OHNMACHT
State University of New York, Albany, NY

READING ABILITY AND MEMORY FOR SENTENCES AND WORDS

Not quite two decades have intervened between the publication of Chomsky's *Syntactic Structures* (1957) and an article by Ortony (1975) entitled "Language isn't for people: On applying theoretical linguistics to practical problems." Although Chomsky's publication was not at all intended to have any direct or obvious relevance for educational issues, Ortony's sharp criticism of most of Chomsky's and his follower's notions is all the more noteworthy because of the influence Chomskyian thinking nonetheless has had on so many language-related issues in education. Not the least of these language-related issues in education, of course, is the area of research on reading processes. Using these two publications as end-points, we intend to indicate in this paper some issues to which we have become more sensitive and in turn which we see as the source of some necessary commitment and investigation. In the broadest sense, we are trying to come to terms with what might be deemed acceptable — or workable — notions of ecological, or contextual, validity (Brunswick, 1956; Bronfenbrenner, 1976) with respect to reading research.

During the last few years some of our own research has reflected attempts to accommodate in our schemes broader contexts than we had heretofore used. In publications of 1972 and 1973 (Fleming, Ohnmacht and Niles; Ohnmacht and Fleming), we varied both deletion strategies as well as contextual constraints in studies of cloze performance. In a 1974 (Ohnmacht and Fleming) paper on developmental changes in memory attributes of good and poor young readers, we found ourselves agreeing with Jenkins' (1974) espousal of a contextualist formulation and acknowledge also that our own studies had yet to exploit in any substantial way children's use of context. The dissertation of one of our students (Judge, 1975) carried further the need for reading researchers more realistically to take advantage of contextual aspects. And in a 1975 (Ohnmacht and Fleming) publication of ours, sentence recognition was studied under varying conditions of semantic and syntactic orienting tasks. Current work with another of our doctoral students (Ohnmacht and Domaracki, in progress) relates reading ability to the recognition of sentences under the conditions of differential orienting tasks just described. To wit, our own work and the work of some of our students increasingly has taken into account a recognition of the importance which must be attached to the knowledge and ability of individuals asked to perform a variety of different reading tasks. These concerns represent a far cry from attempts to characterize, for example, an "ideal" reader — "ideal" in the sense in which

Chomsky (1969) less than ten years ago reaffirmed his goal of describing "competence" in terms of the "ideal speaker-hearer." His notion of "idealization" is striking when set against the enormous amount of research conducted during the past decade or so on just the issues he eschews:

> The idealization is (in particular) that in the study of grammar we abstract away from the many other factors (e.g., memory limitations, distractions, changes of intention in the course of speaking, etc.) that interact with underlying competence to produce actual performance (Chomsky, 1969:12).

Largely as a result of the narrow conditions imposed by Chomsky's insistent "idealization," and despite some revisions in his original formulations, and the rise of competing interpretations given to closely related issues, several rival grammars were to emerge within a relatively short period. Apart from the major role newly assigned to the hitherto relatively neglected area of semantics, there were only a few obvious links between much new thought in linguistics and the increasingly sophisticated investigations of memorial capabilities and processes by psychologists. Both Olson (1970) and Perfetti (1972) comment extensively on, among other matters, the importance given to the role of reference by Fillmore (1968) in the development of his influential notions of case grammar. Tulving and Donaldson (1972) in their edited collection of papers, *Organization of Memory,* present three instances of case-based grammars. A balanced criticism of case grammar as a useful model in research on memory is provided by Reid (1974) who also points up the necessity of keeping in mind the relationship between the choice of informing linguistic theory and the use to which the theory is to be put.

Closely related to the complex and often conflicting theories of reference are the prominent notions of presupposition and inference, both reflecting characteristics of humans processing information with the intent to understand or comprehend. Despite the insights gained as a result of incorporating these notions into even more comprehensive linguistic theories, relatively few attempts have been made to go beyond the level of the sentence in isolation. And as many have noted, individuals normally are not involved in comprehension exercises — oral or written — involving sentences in isolation, although of course individual sentences usually can be understood. Frederiksen (1975b) provides references to the few linguists who have considered the effects of context, including the very few who have attempted to develop text grammars (van Dijk, 1973; Gulstad, 1973). In a related article, Frederiksen (1975a) briefly describes what he maintains is "the most detailed statement of a constructive model of language processing (:) the conceptual dependency system of Schank" (1972).

> In Schank's system, a "conceptual dependency network" is generated from a syntactic parsing of an input sentence *and* from knowledge about the context of an utterance. Constructive models of language processing suggest that comprehension involves using prior knowledge, intentions, context, and task demands in combination with input structure to control discourse processing. Thus, constructive models predict that contexts ought to have pronounced effects on discourse processing (Frederiksen, 1975a: 140).

Interestingly, Schank's work is cited by Norman and Rumelhart (1975) as being extremely influential in their LNR research group which, among other projects, has proposed a sophisticated analysis of reading errors using a network model of grammar, as one product of their over-all network theory. But even some of these comprehensively developed network theories — HAM, for example (Anderson and Bower, 1973) are handicapped by the "static (though not necessarily hierarchical) organization of the knowledge base — a feature of all current network theories" and have come under what seems to be justifiably severe criticism (Anderson and Ortony, 1975). In their insightful paper, "On putting apples into bottles — a problem of polysemy," Anderson and Ortony (1975) present data which cannot be explained by network theories which "assume fixed hierarchical relationships among concepts." As they put one of their major messages in another way: "The extent to which what one knows affects what one understands and stores should not be underestimated" (Anderson and Ortony, 1975: 178).

What we have suggested thus far is that despite the recognition of the need to investigate contextual effects — particularly with regard to the relation between reading comprehension and memory — there has been little help in these matters from linguistic theory and few schemes which allow for investigation beyond the sentence level. In his article, "Language isn't for people," Ortony's (1975) suggestion that many linguists have found "Performance" issues a bit too messy echoes a prominent linguist, Labov (1971) who long has insisted that "a valid linguistic analysis will fit the characteristics of the language used in every-day life when the linguist is not present (Labov, 1971: 415)." When this is not the case, we have, in his words

> . . . the assertion that nothing is to be learned at this stage from "data flux"; that the speech of every-day life [illegible] degenerate data; and that we are not ready to study actual speech (performance) until we have laid a better foundation in the study of competence (Labov, 1971: 438).

Moreover, Labov (1971) concurs with Ortony (1975) with respect to the competence-performance distinction, in that it

> . . . serves to insulate generative grammar from the definition of validity advanced in our first section — that our theories must apply to the unreflecting language used by ordinary people in every-day life (Labov, 1971: 451).

In a later piece, Labov (1973) deals equally as incisively with a major linguistic issue which is relevant to reading research, an issue which also could fit under the "too messy" tag which sometimes is pinned on arguments for greater attention to effects of context. Criticizing semantic feature theories — largely through showing the amorphous conditions surrounding even as concrete an item as a cup — Labov (1973) insists on the need to study variability as a necessary counterpart to the study of invariance. In his words

> The study of variability is the obverse of the study of invariance; one without the other has little significance, and a linguistic study devoted to only one or the other misses the richness of the phenomenon. It is not true that everything varies, anymore than it is true that everything

> remains distinct and discrete. We must locate the boundary between the invariant and variable areas of language with the same precision that we have learned to use in studying the variable elements themselves (Labov, 1973: 367-68).

This point of view could be adopted profitably by many researchers on reading processes, as could Labov's final request "for further research (to) carry us towards experimentation in a more natural setting (Labov, 1973: 369)."

In sum, perhaps more linguists than Ortony (1975) would allow actually are engaged in and lobbying for a more informed, "performance-based" contribution to the resolution of educational issues. In England, Sinclair and Coulthard's (1975) recent work on discourse analysis holds some interesting promise for contextually valid classroom observation studies. So perhaps Ortony is overly pessimistic about the contributions of linguists and his pessimism might be tempered by Baddeley's (1976) "cautious optimism" with regard to "poking our noses outside the laboratory (Baddeley, 1976: 377)" in exploring the psychology of memory.

Beyond the suggestion that contextual factors are of some importance, the varied work of Barclay, Franks and Bransford (cf. Jenkins, 1974) suggests that comprehension of discourse implicates the readers' or hearers' already existing knowledge of the world in that comprehension involves both such knowledge and the nature of the material presented. Such a view points to the importance of knowing considerably more about how knowledge of the world is stored in memory. Models emphasizing semantic features or postulational networks (cf. Collins and Quillian, 1972; Norman and Rumelhart, 1975) may not be adequate to the explanatory roles assigned by some to them since the operators representing presupposition, for example, are not built in. It would seem that theorists are not unmindful of the problem (Munro, 1975: cf in Norman and Rumelhart, 1975). In any event we might suggest that for those interested in the study of reading comprehension that some of the issues we have sketched all too briefly are of substantive import and as has been suggested

> "Hence it is likely that the study of linguistic memory shades imperceptibly into the study of memory in general (Johnson-Laird, 1970)"

So too, the study of reading processes might very much become an aspect of the study of memory in general. To the extent that linguistic theories implicate memory they thus become relevant to the study of reading.

REFERENCES

Anderson, J. R. & Bower, G. H. *Human associative memory.* Washington, D.C.: Winston and Sons, 1973.

Anderson, R. C. & Ortony, A. On putting apples into bottles — a problem of polysemy. *Cognitive Psychology,* 1975, *7,* 167-180.

Baddeley, A. D. *The psychology of memory.* New York: Basic Books, 1976.

Bransford, J. D., Barclay, J. R. & Franks, J. J. Sentence memory: A constructive versus interpretive approach. *Cognitive Psychology,* 1972, *3,* 193-209.

Bronfenbrenner, U. The experimental ecology of education. *Educational Researcher,* 1976, *5,* 5-15.

Brunswik, E. *Perception and the representative design of psychological experiments.* Berkeley, Calif.: University of California Press, 1956.

Chomsky, N. *Syntactic structures.* The Hague: Mounton, 1957.

Chomsky, N. *Topics in the theory of generative grammar.* The Hague: Mouton, 1969.

Collins, A. M. & Quillian, M. R. How to make a language user. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press, 1972.

Fillmore, C. J. The case for case. In E. Bach and R. T. Harms (Eds.), *Universals in linguistic theory.* New York: Holt, Rinehart and Winston, 1968.

Fleming, J. T., Ohnmacht, F. W. & Niles, J. A. *Effects of selected deletion strategies and varying contextual constraint on cloze performance.* Paper presented at National Reading Conference, New Orleans, La., 1972.

Frederiksen, C. H. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology,* 1975a, *7,* 138-166.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. *Cognitive Psychology,* 1975b, *7,* 371-458.

Gulstad, D. E. *A modern theory of 'Langue'.* The Hague: Mouton, 1973.

Jenkins, J. J. Remember that old theory of memory? Well, forget it! *American Psychologist,* 1974, 785-795.

Johnson-Laird, P. N. The perception and memory of sentences. In J. Lyons (Ed.), *New horizons in linguistics.* Baltimore, Md.: Penguin Books, 1970.

Judge, R. E. *The effect of presentation mode and material difficulty on third and seventh graders' use of phonemic and semantic attributes to encode words into long-term memory.* Unpublished doctoral dissertation, The University at Albany, 1975.

Labov, W. Methodology. In W. O. Dingwall (Ed.), *A survey of linguistic science.* College Park, Md.: Linguistics Program, University of Maryland, 1971, 412-497.

Labov, W. The boundaries of words and their meanings. In C-J. N. Bailey and R. W. Shuy (Eds.). *New Ways in Analyzing Variation in English.* Washington, D.C.: Georgetown University Press, 1973, 340-373.

Munro, A. Linguistic theory and the LNR structural representation. In D. A. Norman and D. E. Rumelhart (Eds.). *Explorations in cognition.* San Francisco, Calif.: W. H. Freeman and Company, 1975.

Norman, D. A. & Rumelhart, D. E. (Eds.). *Explorations in cognition.* San Francisco, Calif.: W. H. Freeman and Company, 1975.

Ohnmacht, F. W. & Fleming, J. T. *Some further effects of selected deletion strategies and varying contextual constraint on cloze performance.* Paper presented at National Reading Conference, Houston, Texas, 1973.

Ohnmacht, F. W. & Fleming, J. T. *Developmental changes in memory attributes of good and poor readers.* Paper presented at National Reading Conference, Kansas City, Mo., 1974.

Ohnmacht, F. W. & Fleming, J. T. *Recognition of sentences as a function of semantic and syntactic orienting tasks.* Paper presented at the National Reading Conference, St. Petersburg, Fl., 1975.

Ohnmacht, F. W. & Domaracki, J. H. *The relation of reading ability to the recognition of sentences as a function of semantic and syntactic orienting tasks.* In preparation.

Olson, D. R. Language and thought: aspects of a cognitive theory of semantics. *Psychological Review,* 1970, *77,* 257-273.

Ortony, A. Language isn't for people: on applying theoretical linguistics to practical problems. *Review of Educational Research,* 1975, *45,* 485-504.

Perfetti, C. A. Psychosemantics: some cognitive aspects of structural meaning. *Psychological Bulletin,* 1972, *4,* 241-259.

Reid, L. S. Toward a grammar of the image. *Psychological Bulletin,* 1974, *81,* 319-334.

Shank, R. C. Conceptual dependency: a theory of natural language understanding. *Cognitive Psychology,* 1972, *3,* 552-631.

Sinclair, J. M. & Coulthard, R. M. *Towards an analysis of discourse.* London: Oxford University Press, 1975.

Tulving, E. & Donaldson, W. (Eds). *Organization of memory.* New York: Academic Press, 1972.

van Dijk, T. A. *Some aspects of text grammars.* The Hague: Mouton, 1972.

JACK N. COLE
University of Maryland

THE EFFECTS OF NON-PROSE TEXTUAL CHARACTERISTICS ON RETENTION OF MAJOR CONCEPTS AND SUPPORTING DETAILS WITH AND WITHOUT SPECIFIC INSTRUCTION IN THEIR USE

The use of textbooks in education is nearly universal. However, very little is known about the relative contributions of various format characteristics to the reader's retention of factual content. Past research into the effects of headings (Coles and Foster, 1975; Hershberger and Terry, 1965; Christensen and Stordahl, 1955; and Robinson and Hall, 1941) has yielded results suggesting that headings do not increase retention unless accompanied by instruction. Research into the effects of pictures and graphic representations of information (Peeck, 1974; Davis, 1971; Dwyer, 1970; Frase, 1969; Koenke and Otto, 1969; Magne and Parknas, 1963; and Vernon, 1953) yields results suggesting that, in general, pictures do not enhance retention of textual material. On the other hand, research in the area of memory for pictures themselves (Standing, 1973; Shepard, 1967; and Nickerson, 1965) indicates that humans have an almost unlimited capacity in this regard and that pictures are a highly efficient method of presenting information. Levin holds that pictures may provide certain readers with an "organizational framework" to use in retaining information and, further, that the "training of imagery production in children in need of an organizational framework looms as a reasonable strategy (1976, p. 327)." Although pictures, headings and instructions have all been studied in separate experiments, there is a need to look at them operating in concert.

This study was conducted to discover if retention and reading time could be affected by the presence of headings and illustrations in textbook-like material. This study also investigated whether or not instruction in using these format characteristics would affect retention or reading time.

METHOD

Materials

Passage. The seven-page passage used in this study was a 3258-word selection written by the experimenter and describing three fictitious groups of people living in a fictitious locale. There were four paragraphs for each group of people, describing their origins, physical characteristics, lifestyle and political organization. These paragraphs contained parallel information for each group of people and were structured in the same manner: each paragraph contained a general statement as topic sentence, nine declarative sentences (each stating one detail), and a concluding declarative sentence that served as a transition to the next paragraph. Characteristics attributed to the groups of people were randomly selected from a pool of characteristics. In order to make this selection as much like an actual social studies text as possible, an introductory and a concluding paragraph were written. The SMOG readability procedure (McLaughlin, 1969) estimated the readability of this passage as being at the eighth-grade level.

Four versions of this passage were prepared. The first version contained only the prose text, two paragraphs per page. The second version was identical to the first, except for the addition of headings before the second paragraph on each page. The third version contained illustrations placed before the first paragraph on pages two through seven of the selection.[1] The fourth version contained both headings and illustrations, placed as they were in the second and third versions of the passage.

Instructions. Subjects were assigned to one of two instructional conditions. They were either merely given instructions to read their assigned passages in order to remember the content for later testing or they were given instructions about a specific reading strategy to use in accomplishing this task. These specific strategies corresponded to the four passage versions described in the section above. They were: read and reflect after each paragraph; turn the section heading into a question and then try to answer it after reading the section; look at the picture, predict the content of the following paragraph, and read to confirm the prediction; and a combination of the second and third instructions.

Immediate Recognition Test (IRT). A thirty-item Immediate Recognition Test was used to assess retention of passage content. Three questions were generated for each of the middle ten paragraphs in the selection. One question required information contained in the first sentence (main idea question) and the other two questions required information appearing in sentences two through ten (supporting detail questions). These items were selected from a forty-item test piloted with a sample of subjects as similar as possible to those used in the study in order to assess their passage dependency. The IRT had a coefficient alpha of .86 and a corrected split-half reliability of .94 for the subset of questions selected from those given the pilot group ($N = 36$) and reliabilities of .76 and .78 respectively for the subjects in the main study.

Subjects

Since the literature did not suggest that sex differences existed with respect to the variables being studied, the subjects used in this study were 186 tenth-grade males at a private military school. They were tested during their regularly-scheduled English classes.

Procedures

One hundred eighty-six subjects were randomly assigned to read one of the four passage versions in one of the two instructional conditions. After reading, each subject recorded the time and completed the IRT. Due to incomplete standardized test scores and non-completions of the IRT, the final number of subjects yielding data for analysis was 141. Numbers of subjects receiving general instructions who yielded analyzable data were: for the text only condition, 17; for the text with headings condition, 21; for the text with illustrations condition, 17; and for the text with both headings and illustrations, 15. Numbers of subjects receiving specific reading strategies who yielded analyzable data were: for the text only condition, 17; for the text with headings condition, 17; for the text with illustrations condition, 18; and for the text with both headings and illustrations, 19. Reading times and IRT scores derived from this data collection were analyzed in conjunction with subjects' scores on the National Educational Development Test (Social Studies Reading and Word Usage subtests).

---

[1] Preparation of the illustrations used in this study was done by Ms. Anita Oberright of Kettering, Maryland.

Results

For the eight treatment groups, mean score ranges on the main idea and supporting details subtests were 5.82-6.94 (*SD* = 1.44-2.83) and 14.43-15.15 (*SD* = 2.64-3.56), respectively. Total scores ranged from 21.06-22.07 (*SD* = 3.28-4.85). Two-way analyses of variance of the IRT scores yielded no significant differences between any of the eight treatment groups for main ideas, supporting details or total score. Moreover, Pearson product-moment correlation coefficients computed between IRT scores and NEDT subscores were also nonsignificant.

Two-way analyses of variance for the reading times yielded a significant *F* ratio for the effect of instructions, $F(1,133) = 7.334, p < .008$. Mean reading times for the groups receiving general instructions were: text only, 977 seconds; text with headings, 1027 seconds; text with illustrations, 1045 seconds; and text with headings and illustrations, 1055 seconds. Mean reading times for the groups receiving specific instructions were: text only, 1079 seconds; text with headings, 1108 seconds; text with illustrations, 1146 seconds; and text with headings and illustrations, 1134 seconds. Those subjects receiving instruction in using a specific reading strategy read for an average of 1.5 minutes longer than those merely instructed to read and remember.

DISCUSSION

Illustrations, headings, or a combination of these format characteristics failed to produce differences in recognition scores in this study. Moreover, the slight but systematic differences in reading times associated with individual variations in these characteristics were not statistically significant for any of the eight treatments. These findings concerning the relative contributions of these characteristics operating in concert are in keeping with findings concerning the manner in which they operate separately. It seems as though the subjects in the present study did not utilize these format characteristics to help them retain textual content. This is in keeping with assertions of Robinson and Hall (1941) that college students do not utilize these characteristics to assist them in reading.

Instructions for reading also failed to affect the recognition scores. It is possible that this was due to subjects' initial lack of familiarity with the use of the suggested strategy or that the strategy interfered with the subjects' coding of the information into memory. It is also possible that the mean score on the IRT represents a ceiling on the total information that can be stored by these subjects under these conditions. This possibility seems to be supported by the fact that, though the IRT seems to be passage dependent and adequately reliable, it correlates neither with reading time nor with the NEDT scores.

Instructions increased reading time. However, this effect was significant overall between the group that had been given general instructions and the group that had been given specific reading instructions, but not from one treatment to another. It is possible that more training in the use of the strategies employed would allow results to be assessed in the absence of "first-time" effects. It is also possible that the strategy the subjects were attempting to employ had no effect upon their retention of textbook material. Further research is needed to uncover the factors operating in this situation and affecting memory.

The results of this study indicate a need for further research. Questions raised by the failure of either format characteristics or instructions to boost IRT scores include theoretical as well as practical issues. Not only is research needed to explore the function of long-term memory in situations where students learn from textbooks, but also the function of individual differences and of training in the mnemonic use of these format characteristics need to be systematically examined with a view to improving real-world instruction. The lack of relationship between NEDT scores, reading time and immediate recall may indicate that some memory factor is operating during the processing of textual material in addition to those involved directly in decoding. Possibilities for direct training of such a memory factor are largely unexplored in current educational literature, theory and practice. Further research might yield information contributing to the development of such training procedures.

REFERENCES

Christensen, C. M. & Stordahl, K. E. The effect of organizational aids on comprehension and retention. *Journal of Educational Psychology,* 1955, *46,* 65-74.

Coles, P. & Foster, J. Typographic cueing as an aid to learning from typewritten text. *Programmed Learning and Educational Technology,* 1975, *12,* 102-108.

Davis, O. L. The usefulness of a map with geographic text: a reanalysis of experimental data. *Journal of Geography,* 1971, *70,* 303-306.

Dwyer, F. M. The effect of questions on visual learning. *Perceptual and Motor Skills,* 1970, *30,* 51-54.

Frase, L. T. Tabular and diagrammatic presentation of verbal materials, *Perceptual and Motor Skills,* 1969, *29,* 320-322.

Hershberger, W. A. & Terry, D. F. Typographical cueing in conventional and programed texts. *Journal of Applied Psychology,* 1965, *49,* 55-60.

Koenke, K. & Otto, W. Contribution of pictures to children's comprehension of the main idea in reading. *Psychology in the Schools,* 1969, *6,* 298-302.

Levin, J. R. Comprehending what we read: An outsider looks in. In H. Singer and R. Ruddell (Eds.), *Theoretical Models and Processes of Reading.* Newark, Delaware: International Reading Association, 1976.

Magne, O. & Parknas, L. The learning effects of pictures. *British Journal of Educational Psychology,* 1963, *33,* 265-275.

Nickerson, R. S. Short-term memory for complex meaningful visual configurations: A demonstration of capacity. *Canadian Journal of Psychology,* 1965, *19,* 155-160.

Peeck, J. Retention of pictorial and verbal content of a text with illustrations. *Journal of Educational Psychology,* 1974, *66,* 880-888.

Robinson, F. P. & Hall, P. Studies of higher-level reading abilities. *Journal of Educational Psychology,* 1941, *32,* 241-252.

Shepard, R. N. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior,* 1967, *6,* 156-163.

Standing, L. Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology,* 1973, *25,* 207-222.

Vernon, M. D. The value of pictorial illustration. *British Journal of Educational Psychology,* 1953, *23,* 180-187.

Computer time for the statistical analyses was supplied by the Computer Science Center, University of Maryland at College Park.

E. NOEL BOWLING
West Virginia College of Graduate Studies

JAMES L. LAFFEY
Madison College

CHILDRENS' RETROSPECTIVE ORAL RESPONSES TO SILENT READING TEST ITEMS

Reading researchers have identified some recurring problems related to standardized tests. In fact standardized reading comprehension test items have been subjected to lengthy scrutiny in the past few years.

Smith and Farr (1970), using a high school-college reading test, reported that a large number of items could be answered without the student having read the comprehension passage associated with the test items. Tuinman (1973) in a large scale study of a number of standardized tests used a similar testing strategy and reported that an unusually large number of comprehension test items were not passage dependent.

These two reports clearly suggest that comprehension test items from standardized tests are often passage independent; i.e., readers can accurately answer the comprehension questions without reading the passage associated with them. However, it is important to note that in neither of these studies did the authors attempt to ascertain why the readers were unable to respond accurately to all test items. Furthermore, other than reporting that the readers did respond to specific items accurately, the investigators didn't attempt to ascertain why the students gave the responses to the test items that they did. It is the purpose of this study to describe the reasons why readers respond inaccurately to specific test items. This follows from Goodman's reasoning that all reading responses are caused and that the analysis of inaccurate responses reflects much more than inaccurate word recognition. Goodman (1969) bases his reasoning on the assumption that reading errors, or miscues, are produced through the same processes as those that produce accurate responses. Moreover, through comparing reading errors with the graphic display of the text, inferences can be made regarding the factors that are influencing inaccurate responses.

It should be held in mind, of course, that Goodman was not referring to oral reading as it takes place under the restricted format of standardized testing procedures but rather oral reading which might normally take place in a classroom.

Specifically then, the present study sought to answer the following related questions: (1) Are inaccurate responses to standardized test items due to identifiable factors, or are the responses simply haphazard or capricious? (2) Are individual inaccurate responses idiosyncratic or do particular responses suggest that subjects are using similar strategies in dealing with particular items? (3) Is there evidence to identify error patterns which contribute to inaccurate responses?

METHOD

The population used in this study included two samples. The first sample was comprised of 26 students from a third grade classroom in a rural agricultural region in Virginia's northern Shenandoah Valley. The second sample was made up of nine third-graders from a combination third-fourth grade classroom in a rural mining region of south-central West Virginia. Both samples were selected on the basis of their being readily accessible to the researchers.

Several instruments were used in the study. First, *The Stanford Diagnostic Reading Test,* Level I, Form W, (SDRT) was the test used to collect the data. Secondly, an interview, directed toward specific test items, was used to generate introspective-retrospective responses by the subjects. The interview procedure used followed from Strang's (1967) description of one of 11 techniques available to researchers for studying the reading process. Strang suggests that a researcher can elect to ask students questions about the reading method that they used following the students' reading of a passage. Gibson (1975), too, used a similar technique in reporting the introspective-retrospective responses of mature readers respective to the processes they discerned taking place during their reading.

PROCEDURE

Prior to the administering of the standardized reading test, seven comprehension items (numbers 11, 17, 21, 25, 30, 41 and 42) and nine vocabulary items (numbers 3, 6, 15, 17, 19, 20, 22, 35, and 38) were selected for analysis. During the week following the test administration, 10 students (from Virginia) and all 9 from West Virginia were selected for individual introspective-retrospective interviews. During these interviews, the students were shown individual test items and were asked, "What did you think of when you read this statement?" Each response was then recorded. This procedure was considered to be neutral in that no information was given as to the accuracy of the student's responses to the standardized test items.

A similar procedure was used for encouraging the students to respond to the vocabulary items as was used with the comprehension items. However, due to the test format that requires the examiner to read the test item while the students view only the three answer choices provided, each item of the selected vocabulary items was read to the student. The specific protocol for the vocabulary section of the interview was: "Look at item number *n* on your test. I am going to say what I said when you took the test the other day and I want you to tell me what you thought about when you answered the test item." As with the responses to the comprehension items, the responses and the introspective response data were tabulated and synthesized, using an *N* of 18 for the comprehension test and an *N* of 19 for the vocabulary items.

Analyses

Three steps were used to answer each of the three questions posed earlier.

Step 1. A table of the inaccurate and accurate standardized test responses for both the comprehension and vocabulary sections was prepared to see if unusual response patterns could be detected. A chi-square was computed to see if a particular inaccurate multiple-choice distractor was selected more frequently than what could be anticipated by chance. Furthermore, each explanation related to the inaccurate test response was examined to see if some form of reasoning was apparent during the interview. For the most part, these classifications resulted in mentally exclusive categories.

Step 2. All responses relating to inaccurate standardized test responses were examined on an item by item basis. For example, if from the interviews of six children who responded inaccurately to an item, one particular justification for the inaccurate response was stated by all, then it was assumed that a common rather than a unique factor was influencing this inaccurate response.

Step 3. The groups of responses determined in Step 2 were interpreted and comparisons were made, where possible, *across items.* This was done in order to isolate general factors that may be influencing students' inaccurate responses. For instance, if students were assigning appropriate meanings based on the context but identifying inappropriate distractors, the inappropriate distractors, as well as interview reports, were examined to see if a general strategy could be identified that could explain this pattern of response. Also, comparisons within and between the two sample groups were undertaken to determine if factors influencing a pattern of inaccurate responses were peculiar to one geographical area but not to the other.

RESULTS

The findings are discussed for each question according to the three steps used in the analysis. The comprehension and vocabulary findings are reported separately respective to each question.

Findings related to question one (i.e., Are inaccurate responses to standardized test items due to identifiable factors or are the responses simply haphazard or capricious?) are as follows.

From Table 1 (column Specific Patterns) it can be seen that comprehension items 11, 17, and 25 have too few inaccurate responses to detect any specific patterns. The inaccurate responses for comprehension items 21 and 30 formed some patterns suggesting that the selection of a particular inaccurate distractor was greater than could be anticipated by chance. The explanations given during the interviews substantiate this observation in that only five of the 26 verbal reports contained explanations from which haphazard responses could be inferred. Of these five haphazard responses, three reports indicated that the subjects had difficulty with the format of the test, while two reports indicated that children were guessing in that they were unable to assign meaning to key terms in the stimulus passage.

In short, there were only five cases where students' inaccurate answers were the result of haphazard or capricious responses.

From Table 2 and the data collected during the interviews, it is possible to state that the inaccurate responses were the result of something other than haphazard or capricious responses. Although item three has too few inaccurate

Table 1

Summary of Standardized Scores and
Corresponding Interview
Results Related to Comprehension

	TEST			INTERVIEW				
						3[d]		
Item	Accurate	Inaccurate	Specific Pattern[a]	1[b]	2[c]	A	B	C
11	17	1	—	—	—	—	1	—
17	17	1	—	1	—	—	—	—
21	12	6	*	2	—	4	—	—
25	18	—	—	—	—	—	—	—
30	14	4	*	—	1	3	—	—
41	7	11	—	2	—	3	6	2
42	5	13	—	—	—	—	—	—
Total	90	36	2	5	1	11	7	2

Note. As items 41 and 42 are contained in one stimulus passage, the pupil interview encompassed both test items; tallies sum to 13 (the number of inaccurate responses to item 42).

[a] Denotes an item containing a particular inaccurate distractor that was selected more frequently than what could be anticipated by chance.

[b] haphazard responses

[c] idiosyncratic responses

[d] patterns inferred from interview with A, being responses influenced by students' consideration of related information not contained in the passage; B, responses influenced by students' too limited consideration of information within passage C, responses influenced by decoding, difficulty resulting in inaccurate expectancies.

*$p < .05$ for chi-square test.

responses for determining any specific patterns, some identifiable patterns were present in five of the nine stimulus items. In addition, only three of the ninety-nine verbal reports indicated that the response to the standardized test items were a result of something other than reasoning processes.

Findings related to the second research question (i.e., Are the individual inaccurate responses idiosyncratic or do particular responses suggest that subjects are using similar strategies when dealing with particular items?) are as follows.

First, responses are not idiosyncratic, capricious or haphazard. Tables 1 and 2 indicate that only 13 of a total of 135 responses appeared, after interviewing students, to be idiosyncratic. Comprehension items 21, 30 and 41-42 elicited dominant strategies leading to an inaccurate response, and for items 21 and 30, to a *particular* inaccurate response. Six of the 9 analyzable vocabulary items elicited a dominant strategy leading to an inaccurate response, and, for 4 of those 6 items to a *particular* inaccurate response. One is reminded of Thorndike's (1977) notion of overpotency.

Table 2

Summary of Standardized Scores and Corresponding Interview Results Related to Vocabulary

	TEST			INTERVIEW							
						3[d]					
Item	Accurate	Inaccurate	Specific Pattern[a]	1[b]	2[c]	A	B	C	D	E	F
3	17	2	—	—	—	2	—	—	—	—	—
6	8	11	—	—	—	2	1	—	—	8	—
15	13	6	—	—	—	6	—	—	—	—	—
17	9	10	*	—	1	4	—	5	—	—	—
19	6	13	—	1	—	2	8	—	—	—	2
20	6	13	*	—	—	2	11	—	—	—	—
22	7	12	*	2	3	5	—	—	4	—	—
35	6	13	*	—	—	—	2	—	—	11	—
38	0	19	*	—	—	19	—	—	—	—	—
Total	72	99	5	3	4	42	22	5	4	19	2

[a] Denotes an item containing a particular distractor that was selected more frequently than what could be anticipated by chance.

[b] Denotes haphazard responses.

[c] Denotes idiosyncratic responses.

[d] Denotes error patterns as follows: A = semantic cued response; B = phonological-graphic cued response; C = colloquial cued response; D = graphic cued response; E = dialectical cued response; F = syntactic cued response.

*$p < .05$, for chi-square test.

Findings related to the third research question (i.e., Is there evidence to identify specific error patterns which contribute to inaccurate responses?) are as follows.

As can be seen from Table 1 (columns 3 A, B, C,) three major error groups were observed respective to the comprehension items during the introspective-retrospective interviews. These patterns were abstracted from all interview reports with the exception of those students whose responses were considered haphazard or idiosyncratic.

Error pattern A responses consisted of verbal reports that reflected students misconstruing the information in the stimulus passage due to the student's perception of related experiences. In each of the verbal reports of this pattern, the student justified his response by indicating that his answer was consistent with his knowledge and experiences. For instance, in responding to item 30 (". . . He knew he did not learn to (*skate*) without taking a few *30.* ( ) candies; ( )

friends; ( ) falls; ( ) jackets), the students repeatedly maintained that friends were necessary for learning how to skate.

Error pattern B responses were based on few instances than A. With the B pattern, students tended to disregard parts of relevant information in the passage, this resulting in phenomena commonly called "tunnel vision." The clearest example of this occurred with item 41-42. This passage expressed a comparison between train and truck freight. Interview responses that were categorized as error pattern B ignored the information related to trucks, thus students responded to the comprehension items with information only applicable to trains.

Error pattern C responses were based on only two instances, both occurring for the same test item. With the C pattern, the students inaccurately decoded a distractor in order to make it fit the overall passage. In both instances the students inaccurately decoded "track" or "truck".

As can be noted in Table 2 (Column 3, A-F), 6 major error groups respective to the vocabulary item were abstracted from the interview reports. The student responses which were considered haphazard or idiosyncratic were not included in the following analysis.

It must be held in mind that regardless of the description of a particular error pattern, all error patterns were comprised of reasoned responses and thus were considered as having been semantically influenced.

Error pattern A (Semantic cued responses) were abstracted from verbal reports that reflected students assigning meaning to familiar terms beyond the usual constraints associated with that particular term. The nearest example of this pattern occurred with item 38 (An area of land that is almost surrounded by water is a ( ) swamp ( ) peninsula ( ) nation). In justifying the swamp response, students, without exception, reasoned that swamp was appropriate in that a swamp was land covered by water. Thus it would seem that third graders' conception of the term "surround" exceeds the constraints that would be correctly assigned to that term in that particular context. It can be seen from Table 2 (column 3A) that this error pattern is present respective to all but one of the items.

Error pattern B (phonological-graphic cued response) indicated that students were assigning meaning to terms that sound or appear similar to terms that, if had they been present as a distractor in the test item, would have been correct. An example of error pattern B occurred in response to item 20 (Something that always bothers you is ( ) a patient ( ) an eruption ( ) a nuisance). In justifying the "an eruption" as a response, students clearly indicated that they were responding to this inaccurate distractor as "interruption" rather than "eruption". As can be noted in Table 2 (column 3B), this error pattern is present respective to four of the nine responses.

Error pattern C (colloquial cued-response) were errors in which students were assigning meaning to terms based on the unique way this term is considered in their communities. The only instance of error pattern C occurred in response to item 17 (A large group of trees is a ( ) crop ( ) driveway ( ) grove). Third graders from Virginia justified the response "crop" by making association of this term with the large apple growing industry nearby. Third graders from West

Virginia justified the same response "crop," but for different reasons. In the case of the West Virginian students, "crop" was selected in that it reflected the pulp wood industry nearby.

Error pattern D (graphic cued response) indicated that the students were attempting to derive meaning by matching some visible similarity of a distractor with a key term in the stem of the standardized test item. Although this phenomena may have been influencing the responses noted in pattern B above, it was clearly evidenced in pattern D responses. An example of this pattern can be seen in an answer given to item 22. (When something is usual or common, it is ( ) acceptable ( ) complicated ( ) ordinary). In justifying the response "complicated," one third grader reasoned that complicated looks somewhat like common.

Error pattern E (dialectical cued response) indicated that the students did not differentiate the phonological differences between similar — i.e. pavement-payment and wonder-wander. Error pattern E is similar to Error pattern B except in the case in pattern E responses, students were unaware during the interview that the pairs of terms were different. In contrast to this, students making Error pattern B responses did realize that phonological differences existed between such terms as eruption and interruption.

Error pattern F (syntactic cued responses) indicated that students were predicting a particular term based on a preceding term. This occurred in item 19 (When you try hard, you make ( ) believe ( ) a temptation ( ) an effort). In justifying the believe response, one student stated, "make and believe sounds like they ought to go together."

SUMMARY AND CONCLUSIONS

In summarizing the findings of this study, the authors attempted to answer three research questions related to the inaccuracies of student responses to standardized reading comprehension and vocabulary test items.

In answering question one, there are reasoning factors related to most inaccurate responses made by children to standardized test items. Only in a very few instances did the students reveal that they used a meaningless strategy in responding to the test items.

The answer to question two suggests that there was a great deal of commonality among the students' response strategies. More importantly, their responses during the interviews reveal that the common strategies fall into a discrete number of categories. In only a very few instances were the students' responses idiosyncratic.

Finally, in response to question three, it appears that in responding to both the vocabulary and comprehension items specific error patterns are identifiable.

These conclusions need to be tempered when considering the small number of participants included in the study and the realization that the students were required to rationalize their responses during the interviews. Larger and different samples of test items and students will be required in order to devise a complete and well documented coding scheme. Perhaps the most fitting conclusion is that comprehension behavior, like oral reading behavior (Goodman, 1969), is rational.

REFERENCES

Buros, O. K. (Ed.). *Mental measurement yearbook seventh edition.* Highland Park, New Jersey: Gryphon Press, 1972.

Farr, R. and Smith, C. B. The effects of test item validity in total test reliability and validity. In G. Schick and M. M. May (Eds.), *Reading: process and pedagogy.* Nineteenth Yearbook of the National Reading Conference, Milwaukee: 1970, 122-134.

Gibson, E. J. and Levin, H. *The Psychology of reading.* Cambridge, Mass.: MIT Press, 1975.

Goodman, K. S. Analysis of reading miscues: Applied psycholinguistics. *Reading Research Quarterly,* 1969, *5,* 9-30.

Karlsen, B., Madden, R., and Gardner, E. R. *Stanford diagnostic reading test.* New York: Harcourt, Brace and World, 1966.

Strang, R. Exploration of the reading process. *Reading Research Quarterly,* 1967, *2,* 33-34.

Tuinman, J. J. Determining the passage dependency of comprehension questions in 5 major tests. *Reading Research Quarterly,* 1973, *2,* 206-223.

GARY LEE SHAFFER
James Madison University

AN INVESTIGATION OF THE RELATIONSHIPS OF SELECTED COMPONENTS OF READABILITY AND COMPREHENSION AT THE SECONDARY SCHOOL LEVEL

Most educators agree that the primary goal of schools and the reason for their existence is to promote success for their students engaged in learning activities and academic achievement (Aukerman, 1972); yet, many educators disagree as to how this goal can best be attained. Due to differences in location, populations served, socio-economic conditions, and philosophies it is virtually impossible to find schools any distance apart which are identical in their structure and operation. The one element common to virtually all school systems is the similarity of their curriculum materials. Most curriculum materials are written and require students to be able to read and understand them if they are going to find success while in school.

"One of the problems in public education and mass communication is how to tell whether a particular piece of writing is likely to be readable to a particular group of readers" (Klare, 1974-1975, p. 64). This difficult problem in matching the reader and the material presents the classroom teacher and/or college instructor with a real dilemma when attempting to determine what particular texts, reading assignments, and reading goals will best enable students to reach their academic content goals. Complicating this situation for the teacher and curriculum specialist is the fact that the act of reading is viewed as a very complex process which involves skills at several levels: the abilities of decoding and the abilities involved with acquisition and utilization of meaning which are generally referred to as comprehension (Artley, 1953; Gray, 1948; Russell, 1949; Spache & Spache, 1969; Wilson & Hall, 1972).

Several prominent authors and educators have devised readability formulas in an effort to assess the reading difficulty of written materials and these formulas have been reviewed frequently (Chall, 1958; Gray & Leary, 1935; Klare, 1963, 1974-1975). Most of the readability formulas are mathematical equations based on the length of sentences, the length of words or the number of syllables in words, and the listing of frequently encountered or familiar words chosen from a pre-established word list (Karlin, 1972; Klare, 1974-1975).

To the novice computing these mathematical formulas, it may appear to be a simple matter to "plug" such information from a written passage into one of the many readability formulas available and thus determine the difficulty of a reading passage. However, the process is more complicated than it would appear. As Karlin (1972) indicates, readability formulas provide only a rough measurement of the reading difficulty of materials; grade-placement scores resulting from them should not be taken too seriously since they frequently are too high or too low; and the motivation and interest of readers are ignored.

Readability, then, must be directly related to comprehension if formulas are to be true estimates of reading difficulty. As previously indicated, comprehension is considered to be a vital aspect of reading; therefore, a tie between readability and comprehension must be established. However, authorities in the field of reading cannot agree on what comprehension is or how it can accurately be measured (Spache & Spache, 1969). It is this lack of agreement as to what comprehension is that may be inhibiting the comparison of reading to readability. It is also apparent that the amount of comprehension attained by the reader is related to the difficulty of the passage being read.

The methods used to determine or measure the amount of information obtained from a passage by the reader are many. Among these are informal assessment techniques, cloze procedure, word lists, and readability formulas (Karlin, 1972). In recent years, the cloze procedure has been advocated by several reading authorities as an efficient means of measuring comprehension (Beard, 1967; Bickley, Ellington & Bickley, 1970; Bormuth, 1963, 1966, 1967, 1968; Karlin, 1972; Klare, 1974-1975; Rankin & Culhane, 1969; Schneyer, 1965).

At the present time, researchers are well aware that different written passages vary in their levels of readability. Several studies (Bormuth, 1966; Brown, 1952; Coleman, 1962, 1965; Colvin, 1969; Gibbons, 1941; Hunt, 1965; Klare, Shuford & Nichols, 1957; Nolte, 1937; Peltz, 1973-1974) have been conducted using a variety of variables in an effort to determine how these variables affect readability and comprehension. The results of such studies are inconclusive. As Pearson (1974-1975) states:

> In short, it appears unlikely that any of the correlational or experimental studies in readability has provided a fair test of the variables traditionally assumed to influence comprehension difficulty. The question that must be asked to generate an adequate criterion is, "If I have an idea I want to communicate, what's the best *way* to communicate it?" (p. 162)

Pearson (1974-1975) further states:

> Studies need to be conducted in which different versions of a passage are constructed according to some rule-governed procedures, rather than according to the intuitions of the investigators. (p. 191)

It is the hope of the author that this study will, in some way, meet the request and requirements of not only Pearson but anyone who is interested in the relationship between the components of readability and comprehension.

HYPOTHESES

Hypotheses generated for investigation and testing by this study are as follows:

H_1 There is a significant difference in performance on a cloze test over the same passage by secondary school students from the same grade but with differing reading achievement levels as measured by a standardized test.

$H_{1.1}$ Cloze test performance of students with above average reading ability is significantly greater than cloze test performance of students with average reading ability.

$H_{1.2}$ Cloze test performance of students with above average reading ability is significantly greater than cloze test performance of students with below average reading ability.

$H_{1.3}$ Cloze test performance of students with average reading ability is significantly greater than cloze test performance of students with below average reading ability.

H_2 There is no significant difference in performance on cloze tests by secondary school students with the same reading ability when reading an original passage and when that same passage with its predicted readability level is reduced by: shortening sentences; substituting one or two syllable synonyms for a three-or-more syllable word; substituting synonyms from the *Dale List of 3000 Familiar Words* for words not on the list; and a combination of these factors.

H_3 There is no significant interaction in performance on cloze tests by secondary school students from the same grade with different reading ability levels when reading passages of similar content but different predicted readability levels.

METHOD

Subjects

The population from which the sample ($N = 300$) used in this study was randomly chosen consisted of the sophomore class of the Gulfport Separate School District, Gulfport, Mississippi. Approximately 80% of Gulfport's population is white.

Procedure

Two test instruments, the *California Achievement Tests: Reading*, Level 5, Form A, 1970 edition (CAT:R), and one of five randomly assigned cloze test passages constructed by the author, were administered to each subject during the same testing session. The CAT:R was used as the criterion for identifying above average, average, and below average readers in the sample. Subjects were classified into reading achievement levels by division of raw score percentiles made on the CAT:R into thirds. Cloze test passage scores were used as a criterion to measure amount of comprehension achieved by the sample.

Five different cloze test passage forms were constructed by the author following modified suggested procedures (Karlin, 1972; Taylor, 1953) to measure the amount of comprehension students obtained when reading written passages. In all passage forms, the first and last sentences had no deletions. Counting from the first word in the second sentence, every fifth word thereafter was deleted in all remaining sentences. Subjects were requested to fill-in a blank, 12 typed spaces long, with the word they believed to be deleted. Scoring on all passage forms was by percentage of correct answers with only the deleted word counted as correct when written in the proper blank space. Minor misspellings were not counted as incorrect. Two familiar nouns (radio and television) and one adjective (American) were not altered by the use of synonym substitution in those forms of the cloze test passages in which synonym substitution was a criterion for altering the readability of the passage.

The five cloze test passages used in this study were constructed in the following manner. An original passage (Passage A) was selected from a secondary social studies textbook (Landis, 1964, pp. 184-185) which, through use of the Flesch Reading Ease readability formula, Dale-Chall readability formula, and the Fog Index (Klare, 1974-1975), was determined to have a composite average predicted readability grade level of 10.3, approximately the same as the actual chronological grade placement level of the subjects in the sample.

A second passage (Passage B) was constructed by reducing the length of sentences found in the original passage, taking care to disturb the passage meaning as little as possible. The predicted readability grade level of this passage, determined in the same manner described above, was 8.9.

A third passage (Passage C) was constructed by substituting one or two syllable synonyms, chosen from *Roget's New Pocket Thesaurus in Dictionary Form* (Lewis, 1961), for three-or-more syllable words found in the original passage. The predicted readability grade level of this passage, determined in the same manner described above, was 6.7.

A fourth passage (Passage D) was constructed by substituting synonyms, found on the *Dale List of 3000 Familiar Words*, (Dale & Chall, 1948), for words not found on the list but in the original passage. The predicted readability grade level of this passage, determined in the same manner described above, was 6.9.

A fifth passage (Passage E) was constructed by altering the original passage by all the means described above for the second, third, and fourth passage forms. The predicted readability grade level of this passage, determined in the same manner described above, was 4.1.

Examples of the cloze test passages described above are given below commencing with the first sentence of each passage. Underlined words indicate deletions and the correct response.

Passage A — The machine has recast our life and, with it, patterns of social control. Think of how the *automobile* has modified family life. *Dating* is now in the *automobile* rather than in the *family* parlor.

Passage B — The machine has recast our life. It changed patterns of *social* control. Think of how *the* automobile modified family life. *Dating* is now done in *the* automobile, not in the *family* parlor.

Passage C — The machine has recast our life and, with it, patterns of social control. Think of how the *car* has changed household life. *Dating* is now in the *car* rather than in the *household* parlor.

Passage D — The machine has changed our life and, with it, forms of public checks. Think of how the *automobile* has changed family life. *Dating* is now in the *automobile* rather than in the *family* home.

Passage E — The machine has changed our life. It changed forms of *public* checks. Think how the *car* changed home life. Dating *is* now done in the *car*, not in the home.

Analysis of Data

Two-way analysis of variance (Edwards, 1972) using the AVAR23 computer program (Veldman, 1967) was used to test the previously stated hypotheses. The AVAR23 program allows for traditional testing or analysis of the variance of a set of scores on a dependent variable cross classified by two or three independent variables (Edwards, 1972). In this study, the independent variables were reading achievement level of the subjects and the form of cloze test passage taken by the subjects. The dependent variable was performance on specially constructed cloze test passages. The .05 level of significance was used to test all hypotheses.

RESULTS

Since the analysis of variance revealed no significant interaction, only the means of the main effects are of material value: achievement levels, $F\ (2,285) = 95.45$, $p < .05$, and passages, $F\ (4,285) = 46.95$, $p < .05$. Table 1 shows the computed cell mean scores and the composite main effect mean scores resulting from the administration of the cloze test passages for the sample.

Table 1

Cloze Test Mean Scores for Passages and Achievement Levels

Reading Achievement Level	Passage Form					Mean: Achievement
	A[a]	B[b]	C[c]	D[d]	E[e]	
Above Average	28.388	37.279	30.171	27.625	48.533	34.399
Average	21.104	28.156	18.643	16.608	40.775	25.057
Below Average	14.744	15.877	12.372	9.492	27.155	15.928
Mean: Passage Form	21.412	27.104	20.395	17.908	38.821	

Note. Maximum possible = 100

[a] Predicted readability = 10.3

[b] Predicted readability = 8.9

[c] Predicted readability = 6.7

[d] Predicted readability = 6.9

[e] Predicted readability = 4.1

When a significant *F* was found for the directionally stated hypotheses, Dunn's (1961) Multiple Comparison Test was used to determine where this significant difference lay. Results of this test are shown in Table 2.

Table 2

Dunn's Multiple Comparisons Test* of Difference Between Cloze Test Mean Scores of Above Average, Average, and Below Average Readers

Comparison	Difference	*t*	*p*
Above Average vs. Average	9.3423	6.9880	$< .05$
Above Average vs. Below Average	18.4718	13.8169	$< .05$
Average vs. Below Average	9.1295	6.8288	$< .05$

*$t\,(.05, C = 3, df = 285) = 2.39$

When a significant *F* was found for any of the remaining nondirectional hypotheses, the Scheffé test (Edwards, 1972) was applied to determine where and how the means differed. Results of this test are shown in Table 3.

Examination of the data shown in Tables 1, 2, and 3, resulted in the acceptance of H_1, its three sub-hypotheses, and H_3. Analysis of the data collected revealed that: 1) regardless of the predicted readability level of the passage read, above average readers read cloze test passages significantly better than did average readers who in turn read them significantly better than did below average readers, 2) the cloze test passage E, which was altered from the original by using all methods (readability = 4.1) resulted in significantly better performance than passage B (readability = 8.9), the shortened sentence passage. Both passages E and B resulted in significantly better performance than passage C, the shorter synonym passage (readability = 6.7), or passage D, the Dale List synonym passage (readability = 6.9), or passage A, the original (readability = 10.3). There were no differences among passages C, D, and A.

DISCUSSION

Although the weaknesses inherent in the use of any readability formula have been recognized (Karlin, 1972), they continue to be used extensively (Klare, 1974-1975). If, as is apparently the case, readability formulas are going to continue to be extensively used by classroom teachers and others associated with educational

Table 3
Scheffé Comparisons of Main Effect Mean Scores — Cloze Passages

Comparison of Cloze Passages	Difference	F	p
Original Passage vs. Shorter Sentences Passage	-5.6921	10.92	<.05
Original Passage vs. One or Two Syllable Synonym Substitutions Passage	1.0165	.35	NS
Original Passage vs. Synonym Substitution From the Dale List Passage	3.5037	4.14	NS
Original Passage vs. Combination Passage	- 17.4093	102.14	<.05
Shorter Sentences Passage vs. One or Two Syllable Synonym Substitutions Passage	6.7086	15.17	<.05
Shorter Sentences Passage vs. Synonym Substitution From the Dale List Passage	9.1958	28.50	<.05
Shorter Sentences Passage vs. Combination Passage	- 11.7172	46.27	<.05
One or Two Syllable Synonym Substitutions Passage vs. Synonym Substitution From the Dale List Passage	2.4872	2.08	NS
One or Two Syllable Synonym Substitutions Passage vs. Combination Passage	- 18.4258	114.42	<.05
Synonym Substitution From the Dale List Passage vs. Combination Passage	- 20.9130	147.40	<.05

F' (4,295) at .05 = 9.6

curricular media, it becomes necessary to ask a critical question concerning the assumptions upon which they are based.

The question which must be asked is, "What are the true products of readability?" This can only be established by a direct relationship between students' comprehension of written materials and predicted readability. This study bears on this question.

Readability of a passage is dependent to an extent on the composition of the passage being read, but the reading ability of the person doing the reading is also a factor. There was a direct relationship between performance on cloze tests and the reading achievement level of the reader for all passages, regardless of the predicted readability level of the passage. This was substantiated by the acceptance of the first hypothesis and its sub-hypotheses. Further cloze test performance over all passage forms by all three levels of readers followed the same pattern; i.e., there was no ability across passage interaction.

Not all factors commonly used in readability formulas (i.e., syllabic and common or familiar word counts) affect the degree of comprehension a reader obtains from a written passage to the same degree that they affect predicted readability score (i.e., in a formula). Analysis of variance and Scheffé tests used to test the second hypothesis revealed that no direct relationship existed between the predicted readability level of a passage and the amount of comprehension obtained by subjects when reading these altered passages. Of the five passage forms evaluated, the passage determined by a composite average of readability formulas to have the highest readability rating (original) was shown, in reality, to be read equally as well as the passages with the second (one or two syllable synonym substitutions), and third (synonyms from Dale list) lowest predicted readability rating. Further results from this study show that vocabulary manipulations, by themselves, are not as powerful as sentence length or combination manipulations.

Apparently readability is a global concept which is directly related to comprehension. When all factors evaluated in this study were used to alter the predicted readability of a passage form, that passage was read with greater accuracy and efficiency than when those factors were applied independently to a passage. This was true for all levels of reading ability across all passage forms.

Further research is needed to clarify certain ambiguities in the present study. For example, the boost in cloze test scores that results from a combination manipulation of several variables, each modest or insignificant in its contribution, is remarkable if not inexplicable. It would be useful to look at these manipulations pairwise or in triads. Second, the subtle and not so subtle changes in meaning that result from synonyms changes need to be examined from a semantic point of view. Third, as is necessary in so many comprehension studies, these results should be replicated using other comprehension probes.

While clearcut practical recommendations must await wider sampling of passages, subjects, and reading levels, their study demonstrates a promising methodology for evaluating the effect of textual manipulations on the comprehension of written discourse.

REFERENCES

Artley, A. S. *Your child learns to read.* Chicago: Scott Foresman, 1953.

Aukerman, R. C. *Reading in the secondary school classroom.* New York: McGraw-Hill, 1972.

Beard, J. G. Comprehensibility of high school textbooks: Association with content area. *Journal of Reading*, 1967, *11*, 229-234.

Bickley, A. C., Ellington, B. J., & Bickley, R. T. The cloze procedure: A conspectus. *Journal of Reading Behavior*, 1970, *2*, 232-249.

Bormuth, J. R. Cloze as a measure of readability. In J. A. Figurel (Ed.), *Reading as an intellectual activity.* New York: Scholastic Magazines, 1963.

Bormuth, J. R. Readability: A new approach. *Reading Research Quarterly*, 1966, *1*, 79-132.

Bormuth, J. R. Comparable cloze and multiple-choice comprehension test scores. *Journal of Reading*, 1967, *10*, 291-299.

Bormuth, J. R. Cloze test readability: Criterion reference scores. *Journal of Educational Measurement*, 1968, *5*, 189-196.

Brown, J. I. The Flesch formula. *College English*, 1952, *13*, 393-394.

Chall, J. S. *Readability: An appraisal of research and application.* Columbus: Ohio State University, 1958.

Coleman, E. B. Improving comprehension by shortening sentences. *Journal of Applied Psychology,* 1962, *46,* 131-134.

Coleman, E. B. Learning of prose written in four grammatical transformation. *Journal of Applied Psychology,* 1965, *49,* 332-341.

Colvin, C. M. Psychological principles in materials selection. In J. A. Figurel (Ed.), *Reading and realism.* Newark, Delaware: Proceedings of the International Reading Association Conference, 1969, *13,* 182-185.

Dale, E., & Chall, J. S. A formula for predicting readability: Instructions. *Educational Research Bulletin,* 1948, *27,* 37-54.

Dunn, O. J. Multiple comparisons among means. *Journal of the American Statistical Association,* 1961, *56,* 52-64.

Edwards, A. L. *Experimental design in psychological research* (4th ed.). New York: Holt, Rinehart and Winston, 1972.

Gibbons, H. D. Reading and sentence elements. *Elementary English Review,* 1941, *18,* 42-46.

Gray, W. S. *On their own in reading.* Chicago: Scott Foresman, 1948.

Gray, W. S. & Leary, B. E. *What makes a book readable: An initial study.* Chicago: University of Chicago, 1935.

Hunt, K. *Grammatical structures written at three grade levels.* Champaign, Illinois: National Council of Teachers of English, 1965.

Karlin, R. *Teaching reading in high school* (2nd ed.). Bobbs-Merrill, 1972.

Klare, G. R. *The measurement of readability.* Ames, Iowa: Iowa State University, 1963.

Klare, G. R. Assessing readability. *Reading Research Quarterly,* 1974-1975, *10,* 62-102.

Klare, G. R., Shuford, E. H., & Nichols, W. H. The relationship of style difficulty, practice, and ability to efficiency of reading and to retention. *Journal of Applied Psychology,* 1957, *41,* 222-226.

Landis, P. H. *Sociology.* Boston: Ginn, 1964.

Lewis, N. (Ed.). *Roget's new pocket thesaurus in dictionary form.* New York: Pocket Books, 1961.

Nolte, K. F. Simplification of vocabulary and comprehension in reading. *Elementary English Review,* 1937, *14,* 119-124.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. *Reading Research Quarterly,* 1974-1975, *10,* 155-192.

Peltz, F. K. The effect upon comprehension of repatterning based on students' writing patterns. *Reading Research Quarterly,* 1973-1974, *9,* 603-621.

Rankin, E. F., & Culhane, W. Comparable cloze and multiple-choice comprehension test scores. *Journal of Reading,* 1969, *13,* 193-198.

Russell, D. H. *Children learn to read.* Boston: Ginn, 1949.

Schneyer, J. W. Use of the cloze procedure for improving reading comprehension. *The Reading Teacher,* 1965, *19,* 174-179.

Spache, G. D., & Spache, E. B. *Reading in the elementary school* (2nd ed.). Boston: Allyn and Bacon, 1969.

Taylor, W. L. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly,* 1953, *30,* 415-433.

Veldman, D. J. *Fortran programming for the behavioral sciences.* New York: Holt, Rinehart and Winston, 1967.

Wilson, R. M., & Hall, M. *Reading and the elementary school child: Theory and practice for teachers.* New York: Van Nostrand Reinhold, 1972.

CHESTER E. TILLMAN
University of Florida, Gainesville, FL

READABILITY AND OTHER FACTORS IN COLLEGE READING TESTS: A CRITIQUE OF THE DIAGNOSTIC READING TEST, THE NELSON-DENNY READING TEST, AND THE McGRAW-HILL BASIC READING TEST

Making effective decisions in the process of selecting reading tests for high school and college levels is a continuing problem. Ironically, the increasing popularity of certain tests seems to be inversely related to the negative comments of critics. Farr (1968), for example, after a careful analysis of several reading tests, concluded that the rate subtest of the Nelson-Denny Reading Test (N-DRT) (Brown, 1960), "should not be used at all." Yet the Sweiger survey (1972) found that the number of N-DRT users, 46% of 280 responding junior colleges, was more than double the number using the next most popular test. Undoubtedly, the disappointment expressed by Buros (1972) that test users seemingly have ". . . utmost faith in their own particular choice . . ." of tests, regardless of research or statement of competent critic, applies to those using reading tests.

At the risk of contributing information which seems destined to be ignored, the present author compared the frequently used N-DRT, Form A; the Diagnostic Reading Tests, Survey Section (DRT), Form A (Triggs et al., 1952b), and the less frequently used McGraw-Hill Basic Skills System Reading Test (MGHRT), Form A (Raygor, 1970). Reliability, validity, and readability comparisons among the three tests were based upon data supplied by their respective manuals, an unpublished University of Florida MGHRT intercorrelation study, and a readability analysis of each test. While it is recognized that the three tests might reflect some significant differences in their underlying theories, the comparisons made below are basically empirical.

RELIABILITY

Data in their respective manuals indicate that total score reliabilities for the DRT (KR 21 of .88), the N-DRT (Equivalent Form of .92), and the MGHRT (KR 20 of .89) are within acceptable limits. However, coefficients for each of the three tests should be accepted somewhat cautiously because of speeded subtests. Speeded tests spuriously inflate reliability by forcing retest scores to appear almost identical with initial test scores. Thus, the highly speeded vocabulary sections of the DRT and the N-DRT and the highly speeded skimming section of the MGHRT tend to inflate total reliability. Of particular note is the N-DRT vocabulary section which accounts for 74% of the total score. Also, the reliability of the N-DRT words per minute score (Equivalent Form of .93) which is based upon a one-minute interval would tend to be inflated.

264

VALIDITY

Extensive validity studies reported for the DRT (Triggs et al., 1952a) include intercorrelation studies, correlations with other reading tests, and correlations with ability tests. Notable among the results of seven intercorrelation studies were Vocabulary and Total Score correlations ranging from .919 to .959. Such high intercorrelations raise a question about the importance of the contribution of the comprehension portion of this test. Depending upon the grade levels at which the studies were obtained and the test involved, correlations between the DRT Total Score and the total scores of other reading tests varied from .63 to .79. The DRT Total Score correlations with the verbal score of group general ability tests ranged from .52 to .83, again depending upon the grade level and test involved. Possibly the relatively high vocabulary influence in the DRT accounts for its total score relating to verbal ability tests about as well as it does to other reading ability tests.

Validity data in the N-DRT Form A & B Manual (Brown, 1960) is very inadequate. (More complete data is included in the recently published Form C & D Manual.) The only study reported involves an unspecified number of ninth grade students in a comparison of N-DRT scores with unspecified scores on the "ACE," "Coop English," and the "Otis IQ" tests. Relatively strong N-DRT Total Score correlations with the Coop English ($r = .73$) and the ACE ($r = .83$) probably reflect the heavy vocabulary contribution to the N-DRT Total. Since the "Otis IQ" was not specifically identified, further comment about its correlation of .55 with the N-DRT Total Score cannot be made.

A limited comparison (Raygor, 1970) of MGHRT scores and the N-DRT scores of 67 two-year college students indicated a correlation of .67 between the Total scores of the two tests. Results of two limited intercorrelation studies reported in the manual (Raygor, 1970) are consistent with a more extensive, unpublished correlation analysis involving up to 2,248 University of Florida freshmen. Excluding the flexibility score, which was essentially uncorrelated with any measure, intercorrelations between part scores and Total Score ranged from .31 to .76. Evidently each of the part scores has a separate contribution to make to the total. Correlations with educational achievement as indicated by the Florida Twelfth Grade Testing Program ($r = .59$) and general academic ability as measured by the SCAT Verbal ($r = .50$) and SAT Verbal ($r = .58$) are generally favorable and within ranges expected of most educational variables. Considering that the MGHRT does not have a vocabulary subtest, its substantial relationship with achievement and general ability is evidence that the author has been able to construct a specific and valid measure of reading ability.

Failure of MGHRT Flexibility Score to correlate with other measures suggests the need for further study. Perhaps obtaining a flexibility score by subtracting rate in text materials (Rate 2) from rate in "easy" materials (Rate 1) involves an invalid assumption. What needs to be established is the factorial content of the difference between Rate 1 and Rate 2. Logically, the "flexibility factor" should have a degree of commonality with both Rate 1 and Rate 2. The variance of whatever commonality exists is removed by subtraction, leaving a relatively larger test error and test instruction variance. A relatively larger test error variance in the Flexibility Score would help to account for its lack of correlation with other measures. It is possible that the MGHRT use of the simple difference in rate to indicate flexibility might be an oversimplification of a complex interaction problem.

READABILITY

As with most test manuals, the manuals of these three tests report favorable results of item analyses to indicate difficulty of questions. While it seems that readability of reading test passages should be considered almost as basic as item analyses, the test manuals made little mention of readability estimates. Triggs, et al. (1952a) state that the difficulty level of the DRT rate passage was determined to be seventh or eighth grade level and that the text passages were "scaled by formulae from easy to medium hard." Raygor (1970) states that the Flesch Readability Formula was applied to determine the difficulty level of the two MGHRT reading rate passages but makes no mention of the paragraph comprehension passages. Brown (1960) makes no mention of readability estimates for the N-DRT, Form A and B.

Table 1

Dale-Chall Readability of Selected Diagnostic Reading Test, Nelson-Denny Reading Test, and McGraw-Hill Reading Test Samples

	Intended Grade Range	Instructions to Examinee	Easy Rate Passage	Text Rate Passage	Text Comprehension Passage in Order of Appearance
DRT, Survey Section, Form A		5-6	9-10		7-8, 13-15, 7-8, 9-10[a]
N-DRT, Form A[b]	9-16	9-10		9-10[c]	13-15, 11-12, 16+, 11-12[d]
MGHRT, Form A	11-14	9-10	7-8	16+	7-8, 13-15 16, 16+[e]

Note. [a] All passages evaluated.

[b] According to the manual, for Forms C & D published in 1973, the Flesch Readability Formula was used to select 8th and 9th grade level passages for reading rate. No mention is made of readability for comprehension passages.

[c] "Easy passage" is first of the text paragraphs, no separate comprehension score is given for the rate score.

[d] Passages 2, 4, 6, 8 evaluated.

[e] Passages 1, 2, 3, 4 evaluated; brief interspersed paragraphs not included in the evaluation.

Table 1 summarizes the results of Dale-Chall readability evaluations of Form A of each of the three tests. Instructions to examinees, rate passages, and paragraph comprehension passages were analyzed.

Instructions to Examinees

Overly difficult directions would have an obvious effect upon examinees. Even when the instructions are read along with the examiner as in the three samples evaluated, examinees who tend to score low on standardized tests might be especially discouraged by inappropriate difficulty. Arranging the difficulty of instructions below the intended grade range of the test as was done for the DRT and MGHRT and adjusting interest level appropriately, should allow directions to be easily and inoffensively understood by most examinees.

Rate Passages

Difficulty levels of rate passages on the three tests give greater reason for concern. First, the present 9-10 grade level analysis of the DRT rate passage difficulty is approximately two levels higher than the 7-8 grade level reported by Triggs, et al. (1952a). Perhaps later revisions of the DRT were not subjected to readability evaluation, a formula other than Dale-Chall was used, or different forms of the test vary in difficulty. Since, as Carver (1976) emphasizes, there is very little research interrelating readability, reader ability, and item difficulty, it is not possible to say whether the 9-10 grade difficulty level of the DRT and the N-DRT rate passages are "easy" passages for the grade range intended. The DRT passage appears to be unfairly difficult for seventh and eighth grade examinees. Until more information is available through research, a passage similar in relative difficulty to the MGHRT "Easy" selection, adjusted appropriately for interest, seems to be the most reasonable choice.

Comparing reader performance on the easy rate and text rate passages generates questions about the relative influence of reader flexibility and readability. Normative data reported in the manual and extensive unpublished data accumulated at the University of Florida Reading Center indicate that differences in median easy rate and median text rate have varied from 31 to 54 words per minute. One University of Florida group of 67 Education Transfer Students read the text passage only 45 words per minute slower than the easy passage. Mean comprehension was identical on both passages. Apparently the difference of nearly eight levels in readability made little difference in reading performance for this group. A University of Florida freshman group included in the unpublished data performed more in the expected direction, but differences in rate for the two passages appeared small considering the difference in readability. Such seeming inconsistencies suggest the need for studying the relative limiting effects of reader flexibility and readability.

Paragraph Comprehension Passages

In studying the paragraph comprehension data in Table 1 it is well to remember that the examinee on all three tests is permitted to refer back to any passage to find or check an answer. The three subtests vary considerably in number of passages, in number of questions, and in time limits. While the

MGHRT has a very generous time constraint, the DRT and especially the N-DRT allow relatively little time. Readability data in the table show that the paragraphs of the three tests also vary considerably in difficulty level. On the average the DRT is low, the MGHRT high, and the N-DRT approximately midway. Since college texts vary widely in difficulty level, it might be difficult to say what reading level such a test "should" be. It appears that the DRT passage content is too easy for college students. In addition to average level of passage readability, uniformity of variation in passage readability is also an important consideration. By systematically increasing the difficulty of consecutive passages, the MGHRT arrangement appears to be motivationally sound and fairest for slow examinees. Erratic readability changes from one passage to the next as on the DRT and N-DRT would discourage some examinees and would be unfair to slow readers who might not have an opportunity to attempt easier passages at the end of the subtest. Finally, noticeable in each of the three tests, were sentences extending beyond 85 words in length. The use of such lengthy sentences seems questionable unless it can be shown that they are common in most texts.

By this brief comparison of the three tests it should be evident that readability, by whichever techniques measured, raises questions of concern in reading test construction. Item analyses data can give tests the appearance of being acceptably difficult for the groups offered, even though readability of the reading passages might be unsuitable. Research needs to be undertaken to clarify relationships among reader ability, readability, item difficulty, and percent comprehension.

SUMMARY AND CONCLUSION

Three tests were compared, not especially for the purpose of recommending a particular test, but to re-emphasize the need for caution in selecting tests. Although the MGHRT appears to be the best designed in terms of readability, some weaknesses in reliability, validity, and in readability were noted in each test. Wise use of readability estimates in the construction of these three tests could have probably been used to control motivation and accuracy more effectively. In spite of all their weaknesses, correlations between these three tests and other educational variables are moderate. Indications are that the tests can be useful in prediction, one of their basic purposes. Rather than throw out all tests as some suggest, more care should be exercised in test design, and more constructive interest should be expressed by all professionals in improvement of tests.

REFERENCES

Brown, J. I. *Examiner's manual — the Nelson-Denny reading test, Forms A and B.* Boston: Houghton Mifflin Co., 1960.

Buros, O. K. *The seventh mental measurement yearbook.* Highland Park, N.J.: The Gryphon Press, 1972.

Carver, R. D. Measuring readability and reading ability in order to predict comprehension. In W. D. Miller & G. H. McNinch (Eds.), *Twenty-fifth Yearbook of the National Reading Conference.* Clemson, S.C.: The National Reading Conference, 1976.

Farr, R. The convergent and discriminant validity of several upper level reading tests. In G. B. Schick & M. M. May (Eds.), *Seventeenth Yearbook of the National Reading Conference.* Milwaukee: The National Reading Conference, 1968.

Raygor, A. L. *McGraw-Hill basic skills system reading test — examiner's manual.* Hightstown, N.J.: McGraw-Hill, Inc., 1970.

Sweiger, J. D. Designs and organizational structure of junior and community college reading programs across the country. In F. P. Greene (Ed.), *Twenty-first Yearbook of the National Reading Conference,* Vol. II. Boone, N.C.: The National Reading Conference, 1972.

Triggs, F. O., et al. *Diagnostic reading tests: A history of their construction and validation.* Mountain Home, N.C.: The Committee on Diagnostic Reading Tests, Inc., 1952. (a)

Triggs, F. O., et al. *Diagnostic reading tests, survey section.* Mountain Home, N.C.: The Committee on Diagnostic Reading Tests, Inc., 1952. (b)

ALTON L. RAYGOR
University of Minnesota

THE RAYGOR READABILITY ESTIMATE: A QUICK AND EASY WAY TO DETERMINE DIFFICULTY

Research and development in the area of readability has shown two opposing trends in the past few years. On the one hand, there is the development of multivariate approaches involving very complex computer-based computation of formulas using many variables to achieve a high level of accuracy in prediction of criteria measures (Harris and Jacobson, 1974). On the other hand, there has been some emphasis on development of easier-to-apply formulas that can be used by the teacher who does not have a computer (Fry, 1968).

Many teachers would rather have speed and ease of computations than the extreme accuracy offered by complex formulas (Raygor and Kirsch, 1976). In fact, many teachers will simply do without any evidence of readability rather than go to the trouble of hand calculating a difficult, complex formula. In many cases teachers, reading consultants, clinic staff, and others are called upon (or volunteer) to teach other faculty how to use readability estimates. In such cases they need a formula or graph that is easy to use and understand. The purpose of this article is to present an easy-to-apply approach.

In general, readability formulas use a measure of sentence complexity (e.g. sentence length or number of prepositional phrases) and a measure of word difficulty (word frequency or word length) to derive a regression equation for predicting comprehension difficulty (Klare, 1974-75).

Sentence length has consistently proven to be an accurate and reliable index of sentence complexity. Word difficulty is another matter. When word lists are used to determine, for example, the number of difficult words, the task is laborious without the aid of a computer. The difficulties attached to use of word lists (where each word in a passage must be checked) have led to use of other, easier to determine measures, such as syllable counts, especially by those doing readability estimates without computers.

In addition to the trend toward use of easy-to-determine measures, recent formulas, such as the Fry (1968), have assisted the teacher by presenting graphic aids to computation of a grade level readability score, removing the necessity for arithmetic computation.

The task of applying a typical recent readability formula such as the Fry involves taking 100 word sample passages and doing three things:

1. Determining the number of sentences
2. Counting the syllables in the passages
3. Plotting the two measures on a graph

Anyone familiar with this process is aware that the sentence length count and the plotting of the graph are very quickly done. It is the syllable count that takes most of the time (and introduces the most error). A measure of word difficulty that can be obtained more quickly and accurately would speed up the process a great deal.

Fortunately, such a measure seems to be available. Harris (1974) in discussing the development of several extremely complex computer-based readability formulas which has been shown to be excellent predictors, indicated that one measure of word difficulty that was useful in prediction was the proportion of words of six or more letters. It did not predict as well as word-list variables, but did better than several other measures of word difficulty. Informal experimentation has indicated that one can count the number of words of 6 or more letters in about half the time as one can count the number of syllables in a 100 word passage.

The resulting graph appears at the end of this article. It is used as follows:

1. Count out three 100-word passages at the beginning, middle, and end of a selection or book. Count proper nouns, but not numbers.
2. Count sentences in each passage, estimating to the nearest tenth.
3. Count words with more than five letters.
4. Average the sentence length and word length measures over the three samples and plot the average on the graph. The grade level nearest the spot marked is the best estimate of the difficulty of the selection.

Note that some material will give values that fall far from the line connecting grade averages. Such values suggest that the measure being used is not a valid predictor for such passages, because they represent unusual combinations of word difficulty and sentence length.

DEVELOPMENT OF THE ESTIMATE

The accompanying graph for estimating readability was developed in a fairly standard manner.

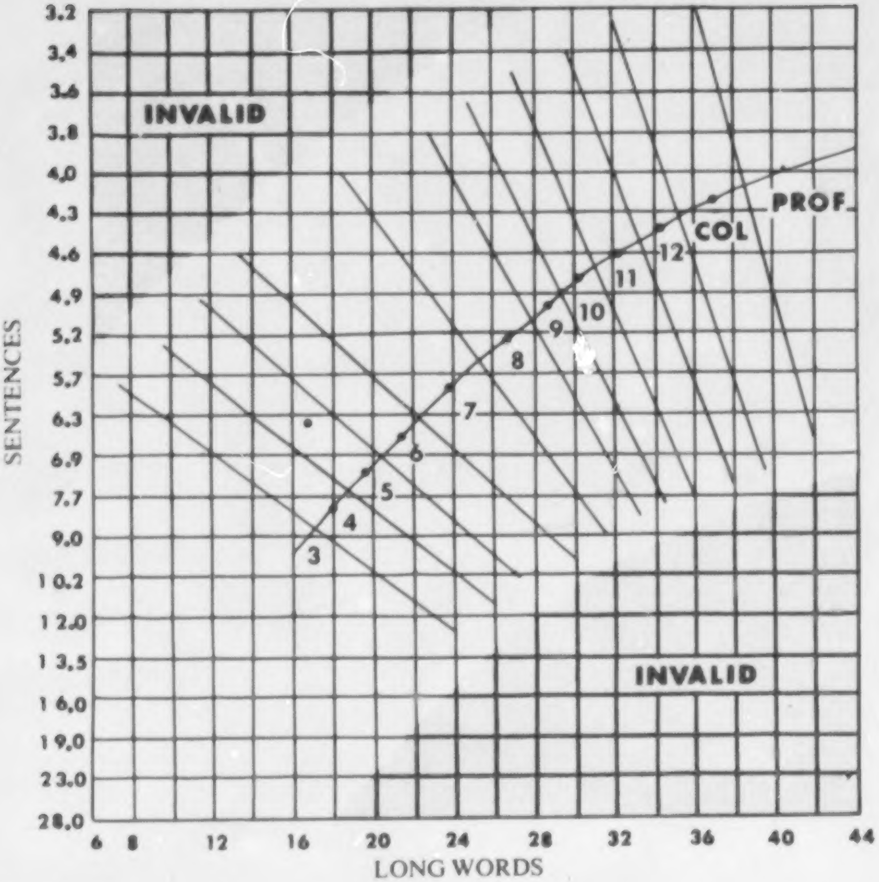
Step 1

Samples of passages were taken from the McCall-Crabbs, *Standard Test Lessons in Reading*, (originally published in the 1920's, revised 1961) children's books, high school tests, and three levels of college texts including samples of vocational-technical school texts, introductory college academic texts, and very advanced, difficult graduate school books.

Step 2

Various readability measures were computed for each passage, and a difficulty level or grade level assigned to each, based on the average of the formulas used. Not all selections were measured using the same formulas. The Spache was not used above sixth grade, for example, and the McCall-Crabbs was not available above the 12th grade. The Spache, the Fry, the Dale-Chall, and the Fog Index (Gunning, 1952) were used on most, in addition to the McCall-Crabbs level.

THE RAYGOR READABILITY ESTIMATE
ALTON L. RAYGOR - UNIVERSITY OF MINNESOTA



Directions:

Count out three 100-word passages at the beginning, middle, and end of a selection or book. Count proper nouns, but not numerals.

1. Count sentences in each passage, estimating to nearest tenth.
2. Count words with six or more letters.
3. Average the sentence length and word length over the three samples and plot the average on the graph.

Example:

	Sentences	6+ Words
A	6.0	15
B	6.8	19
C	6.4	17
Total	19.2	51
Average	6.4	17

Note mark on graph. Grade level is about 5.

Step 3

The two measures, sentence length and word length, were averaged for passages assigned to each grade level, then plotted as points on the graph. The resulting curve was smoothed some, and the cross lines plotted based on deviations.

VALIDATION

It is too early to report as much validity data as would be desired, but generally the results have been encouraging. Correlations with the formulas on which the graph was based are spuriously high, since the original plotting as a mixture of them, so they will not be reported here.

Indirect evidence of validity was available even before the graph was begun, in the original paper presented by Harris (1974). He reported a correlation of .80 between words of more than five letters and basal reader grade level. In addition, a study by Harris and Jacobson (1974) reported use of the sentence length and 5-plus word variable as part of a larger multivariate approach. The two variables were used on a sample of words from six popular basal reader series. They reported a very high level of success with several formulas, the best of which involved computer use of the Harris-Jacobson list. The two variables in the Raygor Readability Estimate, when combined in their formula, gave a Multiple Correlation of .888, with a Standard Error of Estimate of .402 and a Spearman-Brown Reliability of .916. Harris and Jacobson advise use of the word-list based formula, but add that the number of words of more than five letters measure is "faster to obtain and entails only a slight loss in validity."

In summary, the accompanying graph will give a fairly good rough estimate of readability. However, a note of caution is in order. This graph, like any quick and easy method, will give only an approximation (probably within a grade level). It is probably at least as accurate as the Fry, but not as accurate as the Spache, Harris, or Dale-Chall formulas. It is *not* recommended for careful research purposes, for construction or selection of materials in early elementary school, or in any case where a highly reliable and valid measure is needed. For such purposes the author recommends the Spache or the Harris-Jacobson for early grade levels and the Dale-Chall for upper levels. On the other hand, for getting a quick estimate for instructional purposes, it will be reasonably accurate and will save a great deal of time.

REFERENCES

Dale, E., & Chall, J. S. A formula for predicting readability. *Educational Research Bulletin* (Ohio State University), January 21 and February 17, 1948, *27*, 11-20, 37-54.

Fry, E. A readability formula that saves time. *Journal of Reading*, 1968, *11*, 513-516.

Gunning, R. *The technique of clear writing*. New York: McGraw-Hill, 1952.

Harris, A. J. *Some new developments on readability*. Paper presented at Fifth IRA World Congress on Reading, Vienna, Austria, August 13, 1974.

Harris, A. J., & Jacobson, M.D. *Comparative predictions of tested reading comprehension of high school readers*. Paper presented at American Psychological Association Convention, September 1, 1975.

Harris, A. J., & Jacobson, M.D. *Revised Harris-Jacobson readability formulas*. Paper presented at College Reading Association, Bethesda, Maryland, October 31, 1974.

Harris, A. J., & Sipay, E. R. *How to increase reading ability* (6th ed.). New York: David McKay Co., in press.

Klare, G. R. The role of word frequency in readability. *Elementary English*, January 1968, *45*, 12-22.

Klare, G. R. Assessing readability. *Reading Research Quarterly*, 1974-75, *10*, 62-102.

Klare, G. R. *Readability, a new look*. Paper presented at the National Reading Conference, St. Petersburg, Florida, December 1975.

McCall, W. A., & Crabbs, L. M. *Standard test lessons in reading* (Rev. ed.). New York: Teachers College Press, Columbia University, 1961.

Raygor, A. L., & Kirsch, R. Techniques for teaching teachers about readability. *New horizons in reading proceedings of the Fifth IRA World Congress on Reading, Vienna, Austria, August, 1974.* International Reading Association, Newark, Delaware, 1976, 354-356.

Spache, G. D. A new readability formula for primary-grade reading materials. *Elementary School Journal*, March 1953, *53*, 410-413.

Spache, G. D. *Good reading for poor readers* (Rev. ed.). Champaign, Illinois: Garrard Publishing Co., 1974, 195-207.

JAMES A. DINNAN
University of Georgia

PARADIGMATIC/SYNTAGMATIC BEHAVIOR AND THE READING PROCESS

Investigations into word associations have been going on for many decades (Esper, 1973). The most recent investigations have taken a new look at the process by which these associations are being made. No longer are the "products" themselves being the center of attention but rather the "process" as revealed through the intellectual manipulation of the input stimuli. Paradigmatic behavior is the production of a contrast to the stimulus (up-down, high-low). Syntagmatic behavior is predominantly a closure response dependent on personal experience, unique to the individual and outside the expected contrast meanings of general communication (up-tight, high-fly). In terms of the response, whether it be paradigmatic or syntagmatic, it must be considered an intellectual operation. Therefore, each individual reveals an operating intellect and the question must arise as to the difference found in the revealed "products."

Investigations have indicated that when various types of formal testing rank individuals according to their accumulated knowledge, those subjects that perform at or above expected norms produce paradigmatic responses to free association stimuli (Bickley et al. 1970; Rothenberg, 1973). The subjects who perform below level produce predominantly syntagmatic responses. These investigations have used tests such as mental maturity (Cartelli, 1976), reading (Shen and McNinch, 1971; Dinnan et al., 1971a), College Boards (Dinnan et al., 1971b), Graduate Records Exam (Crable, 1975), reading readiness (Lamarre, 1976; Bigaj et al., 1977), and various achievement tests (Brozier, 1974; McNinch, 1972; Baines, 1975) as the independent variable. The dependent variable, subjects' response to words on a free word association test, appears sensitive to all of these types of "achievement" measures.

An analysis of the various "products" produced by the syntagmatic and paradigmatic subjects suggests that syntagmatic subjects respond idiosyncratically (Beauvois, 1973; Dinnan, 1973) while paradigmatic subjects respond normatively; that is, within the expected communication spectrum (Dinnan, 1971).

In order for communication to take place two people must use a common denominator. These common denominators are reflected by paradigmatic behavior and assist the individual to perceive relations within a communication (Dinnan, 1977). More specifically the nucleus of language reflects awareness of time-space-amount relations. It is through these ever recurrent concepts that a human locates the relationship between one matter and another:

time	was	is	will
	had	have	will have
space	from	at	to
	under	on	over
amount	few	some	many
	little	average	big

The common denominator reflected in the "products" produced by a paradigmatically oriented subject indicates that such a subject uses a contrast of the stimulus as the first, immediate, response from the memory (Entwisle, 1966). Thus the base reference falls within the general communication spectrum. The literature indicates that the paradigmatic type associations are not dependent on age but rather are a learned intellectual operation (Rothenberg, 1973; Dinnan and Lodge, 1976).

Training in automatic contrast relations can begin as soon as the child is able to respond to oral stimuli. It can be observed through act or language. This training will insure that when the formal tasks required in schooling begin, the subject will use a common base referent found in language and not a personal closure (syntagmatic operation). Training can be done to correct this serious deficiency in communication at any age level (Cartelli, 1976; Bartell, 1973). Without this intellectual readiness (paradigmatic behavior) the transition from one code, oral language, to another code, written language, will be extremely difficult, if not impossible. Perception is based upon separation, a contrast of one from another. Communication is based upon both the sender and receiver using a common base referent. Human language has a nucleus of references to time, space, and amount (110 words controlling 50% of language redundancy in all human languages), and requires paradigmatic behavior if one is to perceive basic relations. They must be taught in a planned program and not left to chance (Dinnan, 1975).

Investigations into the trainability of paradigmatic behavior have reported success at many school levels. What remains to be done is to convince educators of the importance of developing automatic contrast spectrums as the bases for human understanding and communication.

REFERENCES

Banes, H. *An assessment and comparison of selected linguistic and cognitive abilities of good and poor readers in grades four, eight and twelve.* Unpublished doctoral dissertation, University of Georgia, 1975.

Bartell, H. R., Grill, J. J., & Bartell, H. W. The syntagmatic-paradigmatic shift in learning disabled and normal children. *Journal of Learning Disabilities,* 1973, *6,* 59-64.

Beauvois, J. L. Validity of the P/S criterion for classifying associative responses: Two experiments. *Linguistics,* 1973, *112,* 55-70.

Bickley, A. C., Dinnan, J. A. & Bickley, R. Language responses as a predictor of performance on intelligence tests. *Journal of Reading Behavior,* 1970, *2,* 291-294.

Bickley, A. C., Dinnan, J., & Jones, J. P. Oral associates and reading readiness. In F. P. Greene (Ed.), *Reading: The right to participate.* Twentieth Yearbook of the National Reading Conference, Milwaukee, Wisconsin: The National Reading Conference, Inc., 1971, 14-16.

Bigaj, J. J., Dinnan, J. A. & Crable, E. Word contrasts: Beginning readers' responses. *Journal of Reading Improvement,* 1977.

Brosier, G. *The validation of paradigmatic behavior as relates to selected variables.* Unpublished doctoral dissertation, University of Georgia, 1974.

Cartelli, L. M. *An experimental study in oral language readiness in paradigmatic language behavior in two primary classes of learning disabled children.* Unpublished doctoral dissertation, University of Georgia, 1976.

Crable, E. *A comparison of thought processes as measured by paradigmatic associations, grade point average, graduate record examinations, and faculty ratings, of foreign and native-born graduate students at the University of Georgia.* Unpublished doctoral dissertation, University of Georgia, 1975.

Dinnan, J. A. Key to learning. *Journal of Reading Behavior,* 1971, *3,* 1-13.

Dinnan, J. A. Categorization of paradigmatic-syntagmatic oral responses. *Journal of Reading Behavior,* 1973, *5,* 207-211.

Dinnan, J. A. Intellect "process" used for comprehension of varied discipline areas. *Reading Convention and Inquiry.* Conference Proceedings, Twenty-forty Yearbook, Clemson, S.C.: National Reading Conference, 1975, 247-252.

Dinnan, J. A., Cowart, H. & Bickley, A. C. An analysis of semantic products of disadvantaged and advantaged 2nd, 3rd, and 4th grade good and poor readers. *Journal of Reading Behavior,* 1971, *3,* 22-26.

Dinnan, J., Bickley, R. & Williams, J. Syntagmatic oral responses compared with high and low verbal SAT scores of college freshmen. *Twentieth Yearbook of National Reading Conference,* 1971, 17-19.

Dinnan, J. A., Neilsen, B., & Crable, E. Assessing oral language readiness of slow learners. *Journal of Reading Improvement,* 1976, *3,* 144-146.

Dinnan, J. A. *Communication: A meta theory of language.* Athens, Georgia: Jaddy Enterprises, 1976.

Dinnan, J. A. Questioning the testing of the intellectual process. *Journal of Education,* 1977, *97,* 271-279.

Entwisle, D. R. *Word associations of young children.* Baltimore: John Hopkins University Press, 1966.

Esper, E. A. *Analogy and association in linguistics and psychology.* Athens, Georgia: University of Georgia Press, 1973.

Lamarre, J. L. P. *An analysis of oral language readiness for reading of selected Canadian second-grade students in French and English monolingual and bilingual programs.* Unpublished doctoral dissertation, University of Georgia, 1976.

McNinch, G., Hornsby, J., & Richmond, M. Syntagmatic-paradigmatic shifting in word associations across grade levels. *Southern Journal of Educational Research,* 1970, *4,* 193-206.

McNinch, G. Paradigmatic language training as a method of increasing reading achievement. *Twenty-first yearbook of National Reading Conferences,* Tampa, Florida. 1972, 100-109

Rothenberg, A. Word association and creativity. *Psychological Reports,* 1973, *33,* 3-12.

Shen, S., & McNinch, G. The relationships between language associations and reading achievement in the eighth grade. *Southern Journal Educational Research,* 1971, *5,* 40-50.

LORA M. CARTELLI
University of New Orleans

THE EFFECTS OF PARADIGMATIC LANGUAGE TRAINING ON THE READING PROCESS: AN EXPERIMENTAL STUDY

There has been much research to date (Bickley et al., 1971; Brosier, 1974; Dinnan, 1971; McNinch, 1972) which has consistently shown that there exists a high positive relationship between paradigmatic language behavior (as manifested in free word association tests) and academic achievement (as measured by standardized achievement tests). Children and adults who are paradigmatic in thought content are also academically successful in school. The general conclusions reached from these studies indicate that paradigmatic language structures are inherent in the processes of learning to read and that they are critical behaviors necessary to achievement in learning any written code. There remained the problem of whether paradigmatic behavior could be taught and whether training in paradigmatic language structures would effect reading abilities.

It was the explicit purpose of this study to determine the effects of paradigmatic language training on the reading process with learning disabled children. The main effects and interaction effects of varied independent variables upon criterion variables were investigated. More specifically, this study sought answers to the following questions:

1. What are the effects of special training in paradigmatic language behaviors on reading abilities and paradigmatic responding in children experiencing learning difficulties?
2. How does intellectual ability interact with paradigmatic language training to effect differences in reading abilities and paradigmatic language responding?
3. Are the effects of paradigmatic language training residual or do the effects dissipate after an extended period of time?
4. How do paradigmatic language training, intellectual ability, and treatment with time interact to effect performance on reading tests and paradigmatic language responding?

METHOD

Sample

The sample in this study consisted of children diagnosed as learning disabled ($N = 46$) according to the criteria set forth by the *District Procedures and Guidelines for Special Education* (1975). Out of seven elementary schools, four classes

278

of learning disabled children participated in this study. Two classes (N = 23) served as the experimental group and received paradigmatic language training and two classes (N = 23) served as the control group and received no special training other than that normally obtained in the LD classrooms. The mean chronological age (CA) for the control group was 8.81; SD = 1.03. The mean CA for the experimental group was 8.64; SD = 1.16. The mean IQ score (*Stanford Binet*) for the control group was 104.30; SD = 8.91, and the mean IQ score for the experimental group was 106; SD = 7.72. Mean intelligence test scores were used as the basis for dividing children into high and low ability groups prior to training. Students in the control group receiving scores at or above 104.30 were placed in the high group (N = 12) and those receiving scores below 104.30 were placed in the low group (N = 11). The same procedure was followed the experimental group using a mean of 106. There were 12 students in the high ability group and 11 students in the low ability group.

Materials

Pretests. All LD children in the study were given pre-tests on the *Paradigmatic/Syntagmatic Oral Language Inventory* (Dinnan, 1971) and the *Metropolitan Reading Tests,* Primary II Battery (Durost et al., 1971). The *Stanford Binet Intelligence Scale* (Terman, 1972) was given to those students who did not have this test prior to entry into the LD program. All other children had received the *Stanford Binet* earlier in the 1975 school year. The data from the pre-tests were analyzed according to a one way ANOVA to: 1) determine if differences in initial ability existed between the experimental and control group; 2) insure group equality by statistically equating groups on initial ability if significant differences existed; and, 3) place students in high or low ability groups on the basis of IQ scores. The results of the pre-test findings indicated that no statistically significant differences existed between the experimental and control groups; therefore, they were considered initially equal.

Training Materials. Training in paradigmatic language structures was implemented immediately following the collection of pre-test data. The researcher trained students in each of the experimental classes approximately 2-2½ hours per day for 19 consecutive school days. Total training time consisted of approximately 45-50 hours per class. The *Key to Learning Series,* Kits I and II (Dinnan & Cowart, 1975) was used as the teaching instrument to train students in the acquisition and application of paradagmatic language structures. Students weak in oral paradigmatic behavior were assigned to Kit I, which emphasizes oral language development, before going into Kit II. Students with stronger paradigmatic behavior were assigned initially to Kit II.

Posttest Materials. The first posttesting was conducted immediately following the completion of the experimental training program. Alternate forms of the *P/S Inventory* and the *Metropolitan Reading Tests* were administered to all LD children (both experimental and control classes) at this time. The experimenter waited 10 weeks before conducting follow-up testing. During this interval no LD children received any special training from the experimenter. The purpose of follow-up testing was to determine if training would have any residual effects.

Alternate forms of the same criterion tests that were administered during posttesting were given to all children in the study.

Procedures

The students in each experimental class were divided into two groups within each class (i.e., Kit I and II) for training purposes only; however, training group was not analyzed as a factor in the design. The basic difference between the two training groups centered on students' degree of sophistication in paradigmatic behavior.

The control group (also two classes of LD children) did not receive any special training during this time. They attended their LD programs as usual but were not seen by the experimenter until it was time for the first posttesting.

Data Analyses

To answer the questions posed initially in this study, the data were analyzed according to a factorially-organized repeated measures ANOVA. The main effects and interaction effects of condition (special training vs. no special training), intellectual ability (high group vs. low group), and time of testing (posttest vs. follow-up test) upon paradigmatic language responses and three tests of reading ability (word knowledge, word discrimination and reading comprehension) were tested for statistical significance at the .01 level of confidence.

RESULTS

1. The first question sought to determine the effects of paradigmatic language training on reading abilities and paradigmatic responding. Analysis of the data revealed that LD children who received this special training performed significantly better on all criterion measures than LD children who did not receive special training (Table 1). The *F* (1,42) values

Table 1

Group Means and Standard Deviations
for Dependent Measures

Treatment	Time	P/S Language		Word Knowledge[a]		Word Discrimination[a]		Reading Comprehension[a]	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Experimental (N = 23)	Pre	7.70	6.23	2.07	.70	1.97	.40	1.55	.50
	Post	25.17	2.04	2.60	.84	2.60	.83	2.43	.62
	Follow up	27.65	1.40	3.17	.60	3.14	.85	3.07	.81
Control (N = 23)	Pre	7.39	5.14	2.04	.61	2.01	.67	1.29	.59
	Post	8.05	5.94	2.00	.53	2.11	.61	1.54	.43
	Follow up	9.09	7.28	2.10	.53	2.07	.63	1.59	.51

a = Grade Norm Scores

for all main effects were significant beyond .01: word knowledge, $F = 26.87$, word discrimination, $F = 16.87$, reading comprehension, $F = 61.87$ and paradigmatic language responses, $F = 214.60$.

2. The second question asked by this study sought to determine if special training would be affected by varying degrees of intellectual ability. Do students in the high intelligence group respond better to special training than students in the low group? The results of data revealed that intellectual ability was not statistically significant as a main effect nor as an interaction effect.
3. The third question posed by this study sought to determine if special training in paradigmatic language structures was residual over an extended period of time. The follow-up tests (10 weeks after training ceased) were significantly higher on all reading measures but not for the P/S language test. However, treatment condition by test time interactions for the three reading measures revealed that, for the most part, only children receiving P/S training gained significantly from posttest to follow up test. Inspection of the means in Table 1 (from posttest to follow up test) substantiates this ordinal interaction. Indeed, there do seem to be positive residual effects from P/S language training.

DISCUSSION

The results of this study revealed that paradigmatic language structures are teachable and testable thought products. Inspection of the means and standard deviations for the groups shows that children who received the training made gains which were rapid and positive (note the small *S.D.* for experimental subjects on the P/S posttest and follow-up test). Out of a possible 30 points, the experimental group had already reached a mean score close to the ceiling on the *P/S Inventory*. Further, inspection of the follow-up test means shows that once a paradigmatic language framework was established, it became a part of the cognitive process and did not dissipate over time; the effects of training were found to be residual.

While previous research has shown that paradigmatic association abilities are related to reading achievement, this study indicates that direct intervention in teaching these associations can increase reading skills, at least for the kind of LD children in the sample. According to Underwood (1964), LD children lack, or have not acquired, the memory functions necessary to manipulate input stimuli in an organized fashion. These memory functions may be related to common base of basic linguistic concepts tapped in school assignments and tests and to the concepts taught in the experimental treatment of this study.

If the present data on input organization are in fact representative of the differences between learning disabled and normal learners, then it would appear warranted to: (1) assess childrens' associative processes and (2) implement instructional programs that teach an understanding of the key components of the oral language code and provide for the transfer of this code to reading processes.

REFERENCES

Bickley, R., Bickley, A.D., & Cowart, H. Oral language responses and reading performance in the intermediate grades. In F. P. Greene (Ed.), *Reading: The right to participate.* Twentieth Yearbook of the National Reading Conference, Milwaukee, Wis.: The National Reading Conference, Inc., 1971, 11-13.

Brosier, G. *The validation of paradigmatic behavior as related to selected variables.* Unpublished doctoral dissertation, University of Georgia, Athens, Georgia, 1974.

Dinnan, J. A. Key to learning. *Journal of Reading Behavior,* 1971, *3,* 1-13.

Dinnan, J. A. *The Paradigmatic/Syntagmatic Oral Language Inventory.* Unpublished test, University of Georgia, Athens, Ga., 1971.

Dinna, J. A. & Cowart, H. *The Key to Learning Series,* Kits I and II. In press.

Dinnan, J. A. & Lodge, R. A. *Communication: A meta theory of language.* Athens, Ga.: Jaddy Enterprises, Inc., 1976.

District Procedures for Providing Special Education for EXC: Guidelines. Department of Education for EXC, Division of Elementary and Secondary Education, Bureau of Education, Vol. I, Tallah., Fla., 1975.

Durost, W., Bixler, H., Wrightstone, J., Prescott, G., & Balow, I. *Metropolitan Achievement Tests,* Primary II Battery. New York: Harcourt Brace and Jovanovich, 1971.

Huck, S. & McLean, R. Using a repeated measures ANOVA to analyze the data from a pretest-posttest design: A potentially confusing task. *Psychological Bulletin,* 1975, *82,* 511-518.

McNinch, G., Hornsby, J., & Richmond, M. Syntagmatic-paradigmatic shifting in word association across grade levels. *Southern Journal of Educational Research,* 1970, *4,* 193-206.

Terman, L. & Merrill, M. *Stanford-Binet Intelligence Scale,* Manual for the third revision, Form L-M. Boston: Houghton-Mifflin, Co., 1972.

Underwood, B. The representativeness of rote verbal learning. In A. W. Melton (Ed.), *Categories of human learning.* New York: Academic Press, 1964.

Voyat, G. Thinking before language? A symposium. *Childhood Education.* 1971, *3,* 39-43.

ALBERT J. KINGSTON
JAMES A. DINNAN
KARLYNE KING
The University of Georgia

PARADIGMATIC-SYNTAGMATIC ASSOCIATIONS, READING ACHIEVEMENT AND DOMINANCE IN CULTURALLY DISADVANTAGED SECONDARY STUDENTS

Word association measures which have intrigued psychologists since the days of Wundt and Ebbinghaus have been widely used to study the language and cognitive development of children (Ervin, 1961; Entwisle, et al., 1964; Deese, 1962). More recently researchers probing the relationship of language capability to reading achievement have exploited the technique (Bickley, 1969; McNinch, et al., 1970; Dinnan, et al., 1971a; Dinnan, et al., 1971b). The present study sought to explore the relationship of paradigmatic-syntagmatic responses to handedness-eyedness, reading achievement, and vocabulary.

METHOD

A thirty-word association test designed by Dinnan (1971) was administered to 48 Black, culturally disadvantaged 15 and 16 year olds enrolled in a special reading improvement program. The stimulus words included nouns, adjectives, adverbs, and prepositions and the elicited responses were scored as either being paradigmatic or syntagmatic.* In order to ascertain consistency of type of response, the stimulus word was repeated three times and responses recorded for each occasion. The Nelson Reading Test was administered to ascertain reading grade level. Subjects also were given a simple test to determine eye dominance and asked to demonstrate how they threw a ball to determine hand dominance.

RESULTS

Sixteen pupils were found to be either left-eyed or -handed as compared with 32 who were both right-eyed and right-handed. Reading scores for the group ranged in grade equivalents from a low of pre-primer to a high of 12.9 with a mean of 7.6. The word association measure was scored to determine the number of syntagmatic (sequential) as contrasted to the number of paradigmatic (same word class) responses. Analysis of variance failed to reveal a significant difference between the syntagmatic-paradigmatic response and reaching achievement of subjects who were left-handed or -eyed and those who were right-handed. When the number of syntagmatic responses on the first trial was correlated with

*For those unfamiliar with P/S research see Dinnan's "Paradigmatic/Syntagmatic Behavior and the Reading Process."

Nelson reading scores, however, a significant ($p < .01$) Pearson product-moment coefficient of -.58 was obtained. Analysis of the responses to the word association test indicated that as a whole the group tended to give more syntagmatic and fewer paradigmatic responses than found in a typical high school population. Furthermore, analysis of the three responses indicated that with each successive trial, a greater number of syntagmatic responses was made and the group displayed less language fluency with each successive trial. The mean number of syntagmatic responses for the first trial was 10.6, $SD = 8.39$, for the second 17.7, $SD = 7.05$, and for the third 20.0, $SD = 7.55$.

CONCLUSION

Ervin-Tripp (1973) notes that syntagmatic responses are probable where the variety of contexts following a stimulus word is low relative to its frequency which serves to reduce the number of competing associates. As the child receives formal schooling, there is an increase in the length and variety of sentences employed so that the relative strength of the average syntagmatic association lessens. As the child increases his vocabulary, he learns more contextually similar responses such as antonyms, synonyms, and other so-called abstract words. So, paradigmatic responses are more likely. Students who are more deficient in language are likely to retain more syntagmatic responses longer than those who make normal language development.

In our sample of culturally disadvantaged students, the students as a whole manifested limited awareness or use of semantic congruences, i.e., words of similar and closely related form and class, and displayed greater numbers of syntagmatic responses. The fact that this group on each of the three trials employed a greater number of syntagmatic responses is viewed as evidence of a different progression of language development. This conclusion is substantiated by a mean grade equivalent on the Nelson vocabulary subtest of 7.7 although these students had a minimum of 10 years of schooling. Further, within the sample group itself, those with poorer reading achievements had fewer paradigmatic responses than did those with higher reading ability. Hence, those with higher syntagmatic responses manifested *lower general language facility* than their fellows despite common age, socio-economic status and race as shown by the P/S and Nelson Tests. No relationship between type of response and dominance was found. Those students with high syntagmatic word associations probably would benefit from a program stressing general language improvement of congruent class spectra rather than one with major emphasis on reading skills *per se*. Particularly recommended would be language training designed to develop more precise congruent word class associates. Promising results have been attained by Cartelli (1976).

REFERENCES

Bickley, A. C. *An analysis of frequency, learning rate, and semantic congruence as categorization constraints on beginning readers.* Unpublished doctoral dissertation, University of Georgia, 1969.

Cartelli, L. M. *An experimental study in oral language readiness in paradigmatic language behavior in two primary classes of learning disabled children.* Unpublished doctoral dissertation, University of Georgia, 1976.

Deese, J. Form class and the determinants of association. *Journal of Verbal Learning and Verbal Behavior*, 1962, *1*, 79-84.

Dinnan, J. A., Bickley, R., & Williams, J. C. Syntagmatic oral responses compared with high and low SAT scores of college freshmen. *Twentieth Yearbook of the National Reading Conference*, 1971.

Dinnan, J. A., Bickley, R., & Jones, J. P. Oral associates as an indicator of reading readiness. *Twentieth Yearbook of the National Reading Conference*, 1971.

Entwisle, D., Forsyth, D., & Muus, R. The syntagmatic-paradigmatic shift in children's word associations. *Journal of Verbal Learning and Verbal Behavior*, 1964, *3*, 19-29.

Ervin, S. M. Changes with age in the verbal determinants of word association. *American Journal of Psychology*, 1961, *74*, 361-372.

Ervin-Tripp, S. M. *Language acquisition and communication: Choice essays by S. M. Ervin-Tripp.* Stanford: Stanford University Press, 1973.

McNinch, G., Richmond, M. & Hornsby, J. Syntagmatic-paradigmatic shifting in word associations across grade levels. *Southern Journal of Educational Research*, 1970, *4*, 193-207.

PAUL A. LAMARRE
University of British Columbia

PARADIGMATIC-SYNTAGMATIC RESPONSES OF BILINGUAL STUDENTS AND READING COMPREHENSION

A review of the literature on bilingualism and reading reveals few investigations concerned with the problems of reading in a foreign language. Hatch (1974), in an attempt to determine whether reading should be taught in the native or target language and to determine what factors account for slower reading speed and lower comprehension, reported that the research is confusing; that the questions remain open. She further determined that the slight attention paid to reading in a foreign language is alarming, in light of the numbers of students for whom English is a foreign language.

Large numbers of studies have examined language and academic performance (Hornsby, 1972); language acquisition (Brown, 1973; Cazden, 1972); language background to verbal and non-verbal intelligence (Keats & Keats, 1974), and cognitive consequences of monoglotism and polyglotism (Cummins & Gulutsan, 1974; Tsushima & Hogan, 1975). Others have analyzed the relations between language association and learning (Shen & McNinch, 1972), language and thought (Olson, 1970), associations of children as a function of age, sex, and race (Kahana & Sterneck, 1972), linguistic independence and bilingualism (Macnamara & Kushnir, 1971; Chambers & Jennings, 1975), and the relationship of oral language to reading and other selected variables (Dinnan, 1971).

The examination of cognitive processing with regard to reading performance has led to studies of word association. Several researchers have concluded that, while communication may take place in any language, thought processes are identical. Many maintain that word association is tied to the nature of cognitive processes. Deese (1965) stated: "We study associations in order to make inferences about the nature of human thought, and these associations are cast in the language which embodies the thought" (p. 4).

Some controlled studies reveal definite advantages in cognitive development of bilingual children, refuting a widespread notion that bilingual children have serious cognitive deficits in contrast with their monolingual peers (Kittell, 1959). In view of those programs which showed negative consequences of bilingualism (Tsushima & Hogan, 1975), further research is needed to determine precisely at which point bilingual children encounter difficulty in coping with the global or predominant language of the community, and to explore curriculum design and teaching strategies employed, in order to uncover the constellation of factors which both help and hinder linguistic and reading competence. Chambers and Jennings (1975) found that certain linguistic groups have a greater propensity for functioning within the larger linguistic set (English) than do others. Ianco-Worrall (1972) reported that a high percentage of bilingual youngsters perceived relationships between words symbolically, rather than acoustically.

The purpose of this investigation was to assess the oral language readiness for reading of second-grade students in monolingual and bilingual programs. Test scores of the *Boehm Test of Basic Concepts* (BTBC) and the *Oral Paradigmatic/Syntagmatic Inventory* (P/S Inventory) were the criterion measures used to test for significant differences in the determination of the effects of immersion bilingual (90 percent of the instructional day conducted in one language, 10 percent in another), balanced bilingual (50 percent of the instructional day conducted in one language, 50 percent in the other), and monolingual programs related to reading readiness and free association. The predictor variable was program language.

METHOD

The sample for the study consisted of 220 second grade students enrolled in French monolingual (N = 55), English monolingual (*N* = 55), immersion (*N* = 64), and balanced (*N* = 46) bilingual programs in St. Boniface, Manitoba and Montmagny, Quebec. The P/S Inventory was administered individually to each subject, the BTBC was administered in a group within the classroom setting. All bilingual program group subjects were administered both French and English forms of the instruments employed. Data collection was restricted to the first two weeks of the academic year 1975-1976.

RESULTS

An *F* test and a *t* test for correlated observations were used to compare the variances of the variables. All statistically significant results were subjected to conversions from Tau to Rho statistics in order to determine the magnitude of the effect size of the variables under consideration. The means and standard deviations of specified groups vis-a-vis the dependent variables are presented in Table 1.

Table 1

Means and Standard Deviations by Groups

		Variable			
		BTBC		P/S Inventory	
Program Group	*N*	*M*	*SD*	*M*	*SD*
Monolingual French	55	43.6545	3.3843	7.6545	6.6031
Monolingual English	55	45.7636	2.7622	13.3818	10.0637
Immersion French	64	41.5938	4.5276	6.8438	6.3801
Immersion English	64	45.6563	4.2322	11.9844	9.0440
Balanced French	46	41.5217	5.3986	8.2609	5.5795
Balanced English	46	42.0217	3.7565	11.1087	7.3582
Combined French	110	41.5636	4.8869	7.4364	6.0722
Combined English	110	44.1364	4.4071	11.6182	8.3556

Comparisons were made between monolingual groups, monolingual groups in relation to bilingual groups collectively and independently, and between bilingual groups for each of the instruments used.

With respect to the P/S Inventory, two of the ten analyses produced statistically significant results:

1. Significant differences do exist in paradigmatic language responding between the French and English monolingual group subjects, $F = 12.45$, 1, 108 *df*, $p < .01$. The Eta coefficient is .32, indicating that 10 percent of the variance in paradigmatic language responding is accounted for by program language. Language of instruction does not contribute substantially to its association with the criterion variable.

The current study sought to determine if significant differences existed between the monolingual groups used in this study prior to analysis of the differences between monolingual and bilingual groups. Differences between monolingual groups are rarely considered when the main thrust of the study is bilingual.

2. Significant differences do exist in paradigmatic language responding between the English and French components of bilingual program group subjects, t (108) = 7.35, $p < .01$. English, as the dominant language of the community in which the bilingual respondents live, may account for the significant differences which were found (English $M = 11.6182$; French $M = 7.4364$).

Some of the differences can, perhaps, be explained by the nature of the immersion phenomenon. Many of the subjects have no prior exposure to the immersion language, except incidentally, until it is encountered in the formal school situation. Furthermore, they continue to have no additional exposure to the target language, despite the availability of target language media such as radio, television, and newspapers. If that is truly the case, one year or less of exposure in the immersion language (first grade) can hardly be expected to offset the influence of the native language exposure of seven or more years. English dominance becomes understandable.

DISCUSSION

One of the objectives of this investigation was to ascertain the relationship between reading readiness and paradigmatic language behavior. The orientation of the subjects is basically syntagmatic (according to the criterion for classification of syntagmatic and paradigmatic, means below 15 indicate a syntagmatic orientation), suggesting that they are unable to assimilate much of the foundation work that is presented for mastery in their typical school work. In terms of general expectations in both oral and written language, these subjects, irrespective of program group, do not make associations within the general parameters of language. While it is true that children use the same vocabulary encountered in early reading materials, the associations they make reveal a lack of understanding of relational concepts; concepts which situate key words in relationships that are specific and nonambiguous. If, as Boehm (1971) has suggested in her test manual, these concepts are used, expected, and assumed to be known, and, therefore are rarely, if ever taught, the students must overcome this handicap on their own.

The assumption that usage of prime frequency vocabulary implies understanding of relational concepts and response to meaning has far-reaching, educational implications. If relationships cannot be perceived, failure is imminent. Concomitant with failure is a host of labels which, once applied, become entrenched and difficult to remove.

Therefore, in order to equip children with the basic requirements of language expectancies from which they can build a structure of knowledge, verification procedures must be employed. Association training for students found to be deficient is probably advisable if not necessary. If they are to acquire new concepts on their own, they must acquire specific and automatic contrast responses to basic relational concepts.

A second objective was to determine if bilingualism and monolingualism were related to the ability to respond paradigmatically on a free association test. The non-statistically significant results of the analysis of variance would seem to support a *no-difference* contention. In those instances where there was a statistically significant difference, the magnitude of the effect demonstrated a small relationship. Between 75 and 95 percent of the variance is left unaccounted when program language is employed as the independent variable. Clearly, other factors contribute more substantially to the variance in scores than does program language status.

Because bilingual studies have been concerned with superiority or inferiority of populations, without accounting for the amount of variance attributable to language in the criterion variables, and because bilingual studies have not clearly defined the variables which need to be examined, the equivocation in the literature is justifiable. Method of instruction, comparing those which emphasize close scrutiny of the relational concepts taught, with those that assume such knowledge of relational concepts, is suggested as a more cogent variable to consider.

While it appeared that language of instruction was considered an important variable in oral language readiness for reading, it was not proven to be so. Oral language readiness, using nucleic contrast sets in space, time, and amount, was not established as a potent variable in the present study.

REFERENCES

Boehm, A. E. *Boehm Test of Basic Concepts.* New York: The Psychological Corporation, 1971.

Brown, R. *A First Language: The Early Stages.* Cambridge: Harvard University Press, 1973.

Cazden, C. B. *Child Language and Education.* New York: Holt, Rinehart and Winston, 1972.

Chambers, D. W., & Jennings, S. M. The achievement patterns of eight linguistic sets of children in a pluralistic community. Monograph Number 1 of the Bureau of Research and Field Services, School of Education, University of the Pacific, 1975, *1*.

Cummins, J., & Gulutsan, M. Bilingual education and cognition. *Alberta Journal of Educational Research,* 1974, *20,* 259-269.

Deese, J. *The Structure of Associations in Language and Thought.* Baltimore: Johns Hopkins University Press, 1965.

Dinnan, J. A. The Key to learning. *Journal of Reading Behavior,* 1971, *3,* 1-13.

Hatch, E. Research on reading a second language. *Journal of Reading Behavior,* 1974, *6,* 53-59.

Hornsby, J. The relationship between verbal responses, reading achievement, intelligence, and performance ability for academic and vocational Junior College students. *Southern Journal of Educational Research,* 1975, *9,* 69-82.

Ianco-Worrall, A. D. Bilingualism and cognitive development. *Child Development,* 1972, *43,* 1390-1400.

Kahana, B., & Sterneck, R. Word-association of children as a function of age, sex, and instructions. *Journal of Genetic Psychology*, 1972, *120*, 39-48.

Keats, D. M., & Keats, J. A. The effect of language on concept acquisition in bilingual children. *Journal of Cross-Cultural Psychology*, 1974, *5*, 80-98.

Kittell, J. E. Bilingualism and language-non-language intelligence scores of third-grade children. *Journal of Educational Research*, 1959, *52*, 263-268.

Kittell, J. E. Intelligence test performance of children from bilingual environments. *Elementary School Journal*, 1963, *64*, 76-83.

Macnamara, J., & Kushnir, S. L. Linguistic independence of bilinguals: The input switch. *Journal of Verbal Learning and Verbal Behavior*, 1971, *10*, 480-487.

Olson, D. R. Language and thought: Aspects of a cognitive theory of semantics. *Psychological Review*, 1970, *77*, 257-273.

Shen, S., & McNinch, G. The relationship between language association and reading achievement in eighth grade. *Southern Journal of Educational Research*, 1971, *5*, 40-50.

Tsushima, W. T., & Hogan, T. P. Verbal ability and school achievement of bilingual and monolingual children of different ages. *Journal of Educational Research*, 1975, *45*, 349-353.

SALLY KAMINSKY
The College of Staten Island

ROSE HARRISON
Pace University

MARTHA J. BELL
Brooklyn College

A PRELIMINARY INVESTIGATION OF ORAL LANGUAGE OF BILINGUAL CHILDREN AND THE BEGINNING READING PROCESS

RELATING ORAL LANGUAGE TO BEGINNING READING

This study of language use and language development of first grade bilingual children is a first step in a long-range investigation of oral language production and the beginning reading process. The field of reading has long assumed a relationship between oral language abilities and success in reading, but the exact nature of that relationship has not been explicated. For bilingual children, beginning reading is further complicated by their insufficient control over the language in which reading instruction is given. It can be expected that bilingual children will exhibit varying degrees of ability in Spanish and in English with possible inadequacies in either language for instructional purposes.

The present study is seeking to discover the levels of syntactical complexity and word knowledge that need to be present in bilingual children's oral language in order for them to accomplish the task of learning to read. The investigation attends to the following issues:

1. What levels of language competence do first grade bilingual children possess?
2. How does the bilingual child's language competence relate to successful beginning reading in English and Spanish?
3. What is the appropriate type of reading instruction for bilingual children who may be at various levels of competence in their two languages?
4. What relationship exists between the language used by the classroom teacher and the task-attending behaviors of the children during classroom instruction?

LEARNING TWO LANGUAGES

Although the acquisition of two languages need not result in language confusion and, in fact, may increase mental flexibility and concept formation (Iiams 1976), it is possible that most young bilingual children, whose language is still developing, will have attained only partial knowledge of the grammars of their two languages at the time formal reading instruction begins. There is increasing evidence to support the proposition that learners use strategies in acquiring their second language which are similar to those used in learning their first language (Dulay and Burt, 1974; Buskin, 1976). The learner, however,

291

modifies these strategies during second language acquisition; he has acquired the knowledge that words represent meanings, that these meanings can be extended, assimilated, and differentiated, and that these principles can be applied in learning a new language.

When two language codes are introduced separately, confusion is generally minimized. Overlapping language systems, however, can often produce oversimplification and overextension, particularly at early stages of development when children are testing hypotheses about the language they are using. Since young bilingual children growing up in this country are frequently exposed to two language systems concurrently, it is possible that the second language is introduced at a time when the first language is in a tentative developmental stage. There is reason to believe that beginning school age bilingual children exhibit a wide variety of competencies in the languages they possess ranging along a continuum from monlingualism (in either the natural language or English) to true bilingualism. It is appropriate, therefore, that bilingual education attend to the nature of primary age children's language development and their knowledge of the language in planning instructional programs.

METHOD

Sample

Twenty four first grade children, identified as bilingual in Spanish and in English by their school, were selected as the sample for this study. The children attend a bilingual school which has adopted a plan for teaching reading in both Spanish and English in parallel instructional settings. Each of two first grade classes is paired. A part of each class day is spent with a bilingual teacher who instructs in Spanish only; during the other segment of the day the children are taught in English by a monolingual teacher. Reading instruction is given in both languages.

Language Samples

Examples of language produced by the children have been collected by the researchers prior to formal reading instruction. These early first grade language samples provide a data base from which predictions about beginning reading can be made, and from which the development of language ability can be examined over the entire school year. Three procedures have been used to collect language samples:

1. English and Spanish free speech samples have been analyzed in terms of the grammatical structures children use and the word knowledge they possess. An index of language dominance has been determined based on a comparison of each child's English and Spanisn oral usage.
2. The aural cloze procedure attempts to measure the children's ability to predict oral language sequences in Spanish and in English. This measure is assumed to have some relationship to reading strategies for predicting written language.
3. The Participation-Ratio Observation Scheme measures the children's task-attending behavior during actual classroom instruction and the language

spoken by the children and the teacher. Cohen and Muesser (1975) determined that a relationship exists between child-attending behavior and achievement in reading. The investigators of the present study assume, furthermore, that the language of the classroom teacher has a bearing on the attending behavior of bilingual children.

Procedures

All children in the sample were interviewed by bilingual interviewers in both English and Spanish. These children were presented with picture books without words and asked to formulate their own story. This procedure is modeled after the work of Bloom (1970), Chomsky (1969), and others who used games and toys as stimuli to elicit child speech. Following the free speech sampling, the interviewer employed an aural cloze procedure. Using the picture books without words as a reference for the child, the interviewer deleted words while reading the story aloud. The child was requested to supply the word which best fit the context.

Four stories with deleted words were offered to each child. In two of these selections, deletions occurred every fifth word, as in the standard cloze, so that responses were linguistically randomized. In the remaining two selections, only nouns were deleted in order to elicit knowledge about the child's vocabulary.

The interviews and the tests in Spanish and in English were conducted one week apart for each child. Language order was randomly determined. This procedure attempted to minimize the effect of the language used in the first interview upon the responses elicited in the second interview.

Throughout the school year all children in the sample have been observed for task-attending behavior during Spanish and English instructional periods. The language used by the teacher and that used by the observed children during on-task and off-task activities has been systematically coded. Relationships will be sought between the attending behaviors of children, their language dominance and the language of the instructional setting.

PRELIMINARY RESULTS

Preliminary analyses of the Fall data suggests that Spanish surnamed children from the two first grade classes under study exhibited wide variations in their language knowledge of Spanish and English. At the beginning of first grade, their language production in both languages ranged from control in only one language (English *or* Spanish) to little control in either one. Language control was identified in the study by the extent to which each child produced grammatical sentences and/or complex sentences. Dominance in a language was considered in terms of the child's displaying superior performance in one language over the other language, as well as demonstrating a measure of adequate language control. To be considered bilingual, the child had to exhibit minimum control in both languages.

Of twenty four children interviewed:

(a) about one-third exhibited high control in both languages;

(b) about one-fourth showed medium control in both languages;

293

(c) at least six children were not able to produce sufficient language in either Spanish or English;
(d) some children were clearly dominant in one language (Spanish or English) and knew little or none of the other language used in instruction.

The researchers expect to develop a more detailed analysis of the language structures bilingual children possess so that a beginning model for examining language dominance of young children can be constructed. At present, the evaluation of the children's free speech utterances suggests that the children who possess simple but grammatical sentence structures or those who use more complex structures and verb tenses (with or without grammaticality) are likely to have sufficient control of the languages they speak to respond to reading instruction.

A strong relationship appears to exist between language control and the ability to predict oral language sequences in orally administered cloze tests ($r = .87$). Children were more successful in predicting deletions when only nouns were omitted than when deletions were linguistically randomized. This may be partially due to the tests being aurally received by the children; information after each deletion is not available, as it is in reading. The range of scores on the cloze tests was from 70% to zero for the sum total of four tests, and from 100% to zero on individual tests.

Children with high control in language received high scores on cloze tests; those with low scores were considerably less able to complete the tasks. Language production and language tasks requiring predicting skills appear to be closely related.

FUTURE EFFORTS

Year-long observations of the task behavior of the children in the study will be examined in terms of the language of the classroom and the language the children are using. Final analysis of their language production will be made for the purpose of studying the children's language development during the first grade and the children's first year of formal instruction in reading. Children's reading acquisition in Spanish and in English will be studied for relationships to oral language, task behavior and language used in the classroom. It is expected that a way of describing beginning reading of bilingual children will be evolved so that stages in their reading acquisition may be identified and more adequately understood.

REFERENCES

Bloom, L. *Language development: Form and function in emerging grammars.* Cambridge, Mass.: MIT Press, 1970.

Chomsky, C. *The acquisition of syntax in children 5 to 10.* Cambridge, Mass.: MIT Press, 1969.

Cohen, S. A. & Mueser, A. *How to conduct a p-ratio observation.* Mimeo reprint, 1975.

Guskin, J. T. What the child brings and what the school expects: First and second-language learning and teaching in bilingual-bicultural education. In A. Simoes, Jr. (Ed.), *The bilingual child: Research and analysis of existing educational themes.* New York: Academic press, 1976.

Iiams, T. M. Assessing the scholastic achievement and cognitive development of bilingual and monolingual children. In A. Simoes, Jr. (Ed.), *The bilingual child: Research and analysis of existing educational themes.* New York: Academic Press, 1976.

DIANE SURLE STUPAY
Case Western Reserve University, Cleveland, OH

THE EFFECTS OF LEXICALLY AMBIGUOUS NOUNS EMBEDDED IN A READING TASK FOR CHILDREN

This study was designed to examine children's reading during the on-line processing of connected discourse in order to gain insights into the ways in which meaning operates during the reading process. One way to look more closely at how semantic screens (Goodman, 1973), i.e. the accumulation of meaning from prior context, operate in the reading process is through the study of lexically ambiguous nouns. Lexically ambiguous nouns are defined as nouns that have two or more distinct dictionary meanings. Words such as bat, the animal, and bat, as in baseball, require different contexts and lead the reader to different predictions about context which will follow the ambiguous word.

This study was designed to explore whether the presence of two meanings of a lexically ambiguous noun in contiguous sentences would reduce the size of the unit of the eye-voice span of children at the fourth and sixth grade levels. (The eye-voice span is the number of words the eye is ahead of the voice in oral reading.) Furthermore, since their lexicon is less fully developed (McNeill, 1970; Felzen and Anisfeld, 1970; and Cramer, 1972), it was anticipated that fourth grade children would be less affected by the presence of ambiguity than sixth grade children. Also, the investigation hypothesized that the use of both meanings of the ambiguous nouns would trigger a reprocessing operation, such as Cairns (1973) described, which would be evidenced by miscues in the vicinity of the ambiguous noun.

Past research indicated that decoding proceeds in "chunks" rather than in single words and that these "chunks" are units of meaning (Buswell, 1920; Miller, 1956; Schlesinger, 1968; Wanant, 1968). Smith (1971) noted that while the eye-voice span (EVS) in a skilled reader is approximately four or five words, about the capacity short term memory, it is flexible and can vary from two to six words depending on the structure of the passage being read. Research by Levin and Turner (1968) and Rode (1974-1975) demonstrated that the EVS in children was sensitive to syntactic structures. Recent research by Zollinger (1974) indicated that children's EVS was also sensitive to semantic constraints.

Psycholinguistic research on ambiguity has focused on adult oral language comprehension. In those few studies which did involve reading, the reading process was treated as an analogous comprehension system and was incidental to the research paradigm.

It has been well documented that at some time during sentence processing multiple meanings of lexically ambiguous words are activated in memory

(Lackner and Garrett, 1972; Bever, Garrett and Hurtig, 1973; Foss and Jenkins, 1973; Cairns and Kamerman, 1975). Furthermore, the effects of ambiguity have a greater effect during than after sentence processing. Within the sentence, context prior to an ambiguous item does not appear to bias (i.e. prevent) the activation of both meanings of a lexically ambiguous word (Conrad, 1974; Foss and Jenkins, 1973). Foss and Jenkins (1973) have suggested that in order for the biasing effect of lexical ambiguity "to occur it might be necessary for the biasing word to occur in a clause that is completed prior to the ambiguous word" (p. 586). Cairns (1973) has described a reprocessing operation in working memory when there is a mismatch between information perceived in two sentences. This mismatch and reprocessing would occur when both meanings of a lexically ambiguous noun are presented in contiguous sentences. Therefore, stimulus materials were constructed to tap this reprocessing operation.

This study is one of the few attempts to use a semantic variable, lexical ambiguity, to research the reading process in children. Furthermore, it was designed to extend the knowledge of how lexical screens (Goodman, 1973) operate in a model of reading.

METHOD

Subjects

From a middle-class suburban school system 40 fourth and 40 sixth grade students were chosen randomly from a set of students with average standardized IQ and reading scores as measured by the *California Test of Mental Maturity* and the *Iowa Test of Basic Skills,* respectively. The mean chronological ages of the two grade levels were 10.04 and 11.85 years. The sample was balanced with equal numbers of males and females at each grade level and within each condition.

Apparatus

Thirty-five mm slides with one stimulus paragraph per slide were projected on a Kodak Kodaslide Table Viewer — Model A. The contrast between letters and background on the slides was sufficiently low so as to eliminate any after-image. The eye-voice span was measured by removing the graphic display from view at a point predetermined by the investigator but not known to the reader, and recording the number of words spoken after the graphic display was removed from view.

Stimulus Materials

The stimulus materials used in this study were constructed from a vocabulary list at the third grade level using Harris and Jacobson's *Basic Elementary Reading Vocabularies* (1972). Since difficulty of reading material influences the eye-voice span (Levin and Cohen, 1968), the third grade vocabulary was chosen in order to avoid confounding the results of the EVS measure with reading ability. The stimulus sentences were organized into thirty three-sentence paragraphs, with twelve words in each sentence. One lexically ambiguous noun was embedded in each paragraph. The first sentence in the paragraph was a filler which controlled for silent reading and established context. The second sentence contained one meaning of the ambiguous noun. The third sentence contained a second meaning of the ambiguous noun. Half the time, the ambiguous word appeared in the first noun phrase of either sentence; half the time, in the first verb phrase. For example:

(Filler) STEVE AND JOHN LIKE TO PLAY IN A FIELD NEAR THEIR STREET.

(SENTENCE 1) THE BOYS HAVE A CLUB AND SOME OTHER CHILDREN WANT TO JOIN.

(SENTENCE 2) THEY ALL PLAN A GAME AND A LONG CLUB IS A GUN. Each paragraph contained two possible light-out positions:

1. A one-meaning condition in SENTENCE 1 in which subjects read only one meaning of the ambiguous noun before the graphic display was removed.
2. A two-meanings condition in SENTENCE 2 in which subjects read two meanings of the ambiguous noun, one from SENTENCE 1 and the second in SENTENCE 2.

Within each sentence the light-out position was determined by the location of the ambiguous noun. The light was always turned out immediately prior to the ambiguous word. In SENTENCE 1 above, the light would have been turned out after the subject said the word *A* preceding *CLUB*. In SENTENCE 2 above, the light would have been turned out after the subject said the word *LONG* preceding *CLUB*. There were two possible light-out positions in each sentence: LIGHT-OUT 1 in the first noun phrase, and LIGHT-OUT 2 in the first verb phrase. The number of words following the light-out point to the end of the sentence was the same for each light-out position.

A control condition was established using the same stimulus paragraphs and light-out positions as those in the experimental condition except that in SENTENCE 1 or SENTENCE 2 the ambiguous noun was replaced by an unambiguous synonym, so that the control for SENTENCE 2 of the example given above reads: THEY ALL PLAN A GAME AND A LONG STICK (CLUB) is a GUN. Half of the unambiguous synonyms replaced the ambiguous nouns in SENTENCE 1, and the other half replaced the ambiguous nouns in SENTENCE 2.

Half of the subjects in each condition (experimental and control) saw half of the stimulus material with a light-out position in SENTENCE 1 and the other half of the material with a light-out position in SENTENCE 2. The other half of the subjects saw the same materials in the opposite conditions (light-out in SENTENCE 2 instead in SENTENCE 1).

All paragraphs were randomly ordered for presentation to subjects. Materials were presented to subjects in one of two random orders. Conditions and orders were randomly assigned to subjects.

Procedure

Individually tested subjects were told to read aloud a series of three-line stories on slides. They were further instructed that while they were reading the light would be turned off and they were to "continue reading as far as they could after the light went out." Forty paragraph frames (including ten fillers) were presented one at a time. Filler paragraphs did not contain a lexically ambiguous noun and were used so that subjects would not become sensitive to the presence of ambiguous nouns.

Scoring

Three dependent variables were scored: the EVS, the corrected EVS, and the number of miscues in target SENTENCE 1 and SENTENCE 2 on each paragraph frame. The EVS scoring was based on an errorless count of the exact number of consecutive words a subject reported correctly after the light-out position.

Using the *Goodman Miscue Inventory* (1973), the corrected EVS count differed from the EVS measure in that insertions, substitutions and omissions which were both syntactically and semantically acceptable (i.e. did not disrupt meaning) were used in the EVS count if the subject reported words correctly beyond the miscue. However, the miscued word was not included in the corrected EVS to allow for omissions.

The number of miscues in SENTENCE 1 and SENTENCE 2 were also recorded and verified against a tape of each testing session. Miscues in the filler sentence were not recorded since this research focused on the disruptive effect of lexically ambiguous nouns which did not occur until target SENTENCE 1.

RESULTS

EVS measures, corrected EVS measures, and miscues from the thirty frames read by each subject were averaged by light-out position per subject. A 2^3 factorial analysis of variance technique with repeated measures on two factors (sentence position and light-out position) was performed on each of the three dependent variables. Simple effects t tests were then run since we were analyzing a priori comparisons (Winer, 1971, p. 634).

Eye-Voice Span Measures

The analysis of variance for the EVS did not result in a main effect finding for experimental-control conditions. However, the analysis of variance for the EVS was complicated by a significant 3 way interaction between experimental-control conditions, grade level and light-out positions, $F(1, 72) = 7.27, p < .01$. The results of two-tailed t tests comparing experimental and control EVS means at each grade level for each light-out position are represented in Table 1.

Table 1

Simple Effects t Tests for EVS Interaction
Comparing Experimental-Control Conditions
by Grade Level and Light-Out Positions

Grade Level	Light-Out	Means		t_{obs}
		Control	Experimental	
4	1	2.309	1.809	2.588*
4	2	2.488	2.379	.5798
6	1	2.818	2.773	.233
6	2	3.040	3.130	.466

*$p < .05$.

Since there was a significant difference between experimental and control means for fourth grade students when the light went out during the 1st noun phrase, these data provided support for the hypothesis that the presence of both meanings of the ambiguous nouns would increase processing time and therefore decrease the EVS. Sixth grade subjects appeared to have enough control of the reading process (or perhaps enough sensitivity to context) to compensate for the disruptive effects of lexical ambiguity as measured by an EVS count.

Corrected Eye-Voice Span Measures

An analysis of variance performed on the corrected EVS scores resulted in a significant main effect for the experimental-control variable, $F(1, 72) = 4.02$, $p < .05$, with the control condition ($M = 2.732$) yielding a longer corrected EVS than the experimental condition ($M = 2.533$). Apparently while the difference between the control means was relatively small, there was almost no difference in the means under the experimental condition. Possibly this would indicate that the extra time required to process both meanings of the ambiguous noun reduced the opportunity to reprocess miscues, since reprocessing inserted an extra step into working memory (Cairns, 1973).

As in the EVS measure, a significant 3 way interaction between experimental-control conditions, grade level and light-out position $F(1, 72) = 7.5$, $p < .01$, complicated interpretation of the main effects. Two-tailed t tests comparing experimental-control means at each grade level for each light-out position revealed that the only significant difference between experimental and control corrected EVS means was for fourth grade subjects at LIGHT-OUT 1 ($p < .05$, see Table 2).

Table 2

T Tests on Corrected EVS Interaction Comparing
Experimental-Control Conditions by
Grade Level and Light-Out Position

Grade Level	Light-Out	Means		t_{obs}
		Control	Experimental	
4	1	2.334	1.829	2.541*
4	2	2.563	2.403	.809
6	1	2.813	2.828	.015
6	2	3.220	3.074	.736

*$p < .05$.

Again, it was the fourth grade subjects who were significantly affected by the presence of two meanings of the ambiguous nouns.

The analysis of variance on both the EVS and the corrected EVS count resulted in a significant simple effect for experimental subjects for light-out position in grade four, with the shorter EVS at LIGHT-OUT 1, the first noun phrase. This finding seems to indicate that for fourth grade students, less contextual input results in shorter EVS when ambiguous nouns are present.

The corrected EVS analysis of variance also revealed a significant finding for the main effect of sentence position. The mean for SENTENCE 1 was 2.699 compared to a mean of 2.567 for SENTENCE 2. On the surface, this larger mean for SENTENCE 1 seems to conflict with the interpretation of the main effects finding for light-out position. However, a significant grade level by sentence position interaction resulted in a higher mean on SENTENCE 1 at the sixth grade level, t (72) = 2.99, $p > .01$, but not at the fourth grade level. While the sixth grade subjects were not differentially affected in the experimental-control condition, it appeared that the shift in context in the second sentence, as well as the presence of ambiguous nouns, reduced their corrected EVS measures.

Miscues per Light-Out Position

The analysis of variance on miscues per light-out position resulted in a significant four-way experimental-control by sex, by sentence position, by light-out position interaction, F (1, 72) = 5.92, $p < .05$. To verify the source of significance for this interaction t tests were done comparing boys and girls on the experimental-control condition at each sentence level and at each light-out position. The difference between means in SENTENCE 2, LIGHT-OUT 1 was significant for the girls in the experimental condition, $^{t}\mathrm{obs}$ (72) = 2.36, $p < .05$. Also, the girls made significantly fewer miscues than the boys in SENTENCE 2, LIGHT-OUT 1 in the control condition, $^{t}\mathrm{obs}$ (72) = 2.43, $p < .05$. Across grade levels the girls were more affected by the presence of both meanings of the ambiguous nouns (i.e. in the experimental condition) as evidenced by a significantly larger number of miscues in this condition. This finding provided partial support for the hypothesis that the use of two meanings of the ambiguous nouns would result in a reprocessing operation which would be evidenced by a higher number of miscues. Unfortunately, this was found only for females.

DISCUSSION

The major question assessed in this research was whether the presence of both meanings of lexically ambiguous nouns in contiguous sentences would increase processing time and therefore reduce the EVS. The data summarized here are largely in accord with this supposition. However, it was the fourth grade subjects rather than the sixth grade subjects who were significantly affected in the experimental condition. It appeared that in order to measure the reprocessing effects of ambiguous nouns, the reading task needed to be moderately difficult. In an experiment designed to evaluate dependent variables that measured comprehension, Hakes and Cairns (1970) found that in order to assess comprehension, stimulus materials needed to be difficult enough to yield a moderate to high rate of errors.

Possibly, these materials were sufficiently difficult for fourth, but not for sixth, grade subjects to provide a reasonable index of their ability to comprehend written text. It is also possible that similar stimulus materials rewritten at a fifth grade reading level would reveal the reprocessing of ambiguous information in sixth grade subjects.

The sex difference finding on the miscues per light-out position measure was unexpected. It was unlikely that the girls in this study were better readers than the boys since both sexes met the same criteria on standardized reading scores. Maccoby and Jacklin (1974) found that studies of reading ability and verbal skills did not show consistent sex differences. However, the authors noted that there appeared to be "distinct phases in the development of verbal skills in the two sexes through the growth cycle;" at about ages ten and eleven girls "come into their own in verbal performance" (p. 84). The present study seemed to have tapped the pre-adolescent developmental spurt; the average age of the subjects was 10.04 and 11.85. This added verbal power appeared to give the girls in this study somewhat of an advantage in using semantic information in the reading process.

The major implication of this study is that semantic information is an integral part of the reading process and that even relatively young readers, ten year-olds, are sensitive to the disruptive effects of lexical ambiguity. It appears that the semantic lexicon of ten year-olds is fully sensitive to the semantic constraints of a third grade vocabulary. Furthermore, this study provides empirical evidence that the semantic screens hypothesized in Goodman's (1973) model of reading do in fact operate in the reading process and that pre-adolescent girls may be somewhat more sensitive than boys to semantic constraints in the reading process.

REFERENCES

Abrams, S. G., & Zuber, B. L. Some temporal characteristics of information processing during reading. *Reading Research Quarterly,* Fall 1972, *1,* 40-51.

Bever, T. G., Garrett, M. G., & Hurtig, R. The interaction of perceptual processes and ambiguous sentences. *Memory and Cognition,* 1973, *1,* 277-286.

Buswell, G. T. An experimental study of eye-voice span in reading. *Supplementary Educational Monographs* (No. 17). Chicago: University of Chicago Department of Education, 1920.

Cairns, H., & Kamerman, J. Lexical information processing during sentence comprehension. *Journal of Verbal Learning and Verbal Behavior,* 1975, *14,* 170-179.

Cairns, H. Effects of bias for processing and reprocessing of lexically ambiguous sentences. *Journal of Experimental Psychology,* 1973, *79,* 337-343.

Conrad, C. Context effects in sentence comprehension: A study of the subjective lexicon. *Memory and Cognition,* 1974, *2,* 130-138.

Cramer, P. A. A developmental study of errors in memory. *Developmental Psychology,* 1972, *F,* 204-209.

Felzen, E., & Anisfeld, M. A. Semantic and phonetic relations in the false recognition of words by third and sixth grade children. *Developmental Psychology,* 1970, *3,* 163-168.

Foss, D., & Jenkins, C. Some effects of context on the comprehension of ambiguous sentences. *Journal of Verbal Learning and Verbal Behavior,* 1937, 577-589.

Goodman, K. *Theoretically based studies of patterns of miscues in oral reading performances* (HEW Project No. 9-0375, Grant No. OEG-0-9-320875-4269). Washington, D. C.: U.S. Department of Health, Education and Welfare, 1973.

Harris, A., & Jacobson, M. *Basic elementary reading vocabularies.* New York: The Macmillan Company, 1972.

Hakes, D. T., & Cairns, H. Sentence comprehension and relative pronouns. *Perception and Psychophysics,* 1970, *8,* 5-8.

Lackner, J., & Garrett, M. Resolving ambiguity: Effects of biasing context in the unattended ear. *Cognition,* 1972, 359-372.

Levin, H., & Cohen, J. Effects of instruction on the eye-voice span. In H. Levin, E. Gibson, & J. J. Gibson (Eds.), *The analysis of reading skills* (Final Report, Project No. 5-1213). From Cornell University to U.S. Office of Education, December 1968.

Levin, H., & Turner, A. Sentence structure and eye-voice span. In H. Levin, E. J. Gibson, & J. J. Gibson (Eds.), *The analysis of reading skills* (Final Report, Project No. 5-1213). From Cornell University to U.S. Office of Education, December 1968.

Maccoby, E. E., & Jacklin, C. N. *The psychology of sex differences.* Stanford, California: Stanford University Press, 1974.

McNeill, D. A. *The acquisition of language.* New York: Harper and Row, 1970.

Miller, G. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81-97.

Rode, S. Development of phrase and clause boundary reading in children. *Reading Research Quarterly*, 1974-1975, *10*, 124-142.

Schlesinger, I. M. *Sentence structure and the reading process.* The Hague: Mouton, 1968.

Smith, F. *Understanding reading: A psycholinguistic analysis of reading and learning to read.* New York: Holt Rinehart and Winston, Inc., 1971.

Wanant, S., & Levin, H. The eye-voice span: Reading efficiency and syntactic predictability. In H. Levin, E. Gibson, & J. J. Gibson (Eds.), *The analysis of reading skills* (Final Report, Project No. 5-1213). From Cornell University to U.S. Office of Education, December 1968.

Winer, J. B. *Statistical principles in experimental design.* New York: McGraw-Hill, 1971.

Zollinger, R. H. *The psychological reality of information focus for the reader.* Unpublished doctoral dissertation, Case Western Reserve University, 1974.

SEAN A. WALMSLEY
State University of New York at Albany

CHILDREN'S UNDERSTANDING OF LINGUISTIC CONNECTIVES: A REVIEW OF SELECTED LITERATURE AND IMPLICATIONS FOR READING RESEARCH

Research on children's understanding of logical connectives has provided some important evidence of the nature and role of such words as *and, or* and *because* in speaking and comprehending connected discourse (Piaget, 1968; Lakoff, 1971): since these connectives also are related to the propositional logical relationships believed to underlie the sentences in which they appear, research in this area contributes to the broader issue of language and thought (Vygotsky, 1962).

A linguistic connective (or logical or language connective — the terms appear to be used according to the orientation of a writer's discipline) may be defined as a "co-ordinating, qualifying or adverbial conjunction used to link a simple proposition with another idea (either a proposition or a concept) to form a complex proposition" (Gardner, 1975). Alternatively, it may be defined as a syntactic structure signalling underlying logicosemantic relations (Olds, 1968). Connectives may link propositions within or between sentences; they may comprise a single word (e.g. *and*), or a phrase (e.g. *in addition to*). A single connective may signal quite different underlying logicosemantic relations.

Brief review of research on linguistic connectives

(a) *General perspectives.* Most of the research on linguistic connectives has focused on their role as syntactic signals for specifying underlying propositional relationships (Neimark & Slotnick, 1970; Paris, 1973; Bart, 1974; Beilin & Lust, 1975). Three connectives in particular have been the focus of attention — *and,* expressing the logical relationship of *conjunction* (Paris, 1973; Beilin & Lust, 1975); *or,* signifying *disjunction* (Neimark, 1970); and *if-then,* denoting conditionality (Olds, 1968; Shapiro & O'Brien, 1970).[1]

Some authors have attempted, using cross-sectional data, to show how children's use and understanding of linguistic connectives undergoes developmental changes between ages 4 and adulthood (Neimark & Slotnick, 1970; Beilin & Lust, 1975). Others have focused on developmental changes over shorter time periods in children's use and interpretations of particular connectives. For example, Piaget shows how children come to learn two 'meanings' for *because* — the relation of cause and effect, and the relation of reason and consequent (Piaget, 1968; p. 6).

[1] Conjunction and disjunction are terms from propositional calculus, equivalent to set intersection and set union in Boolean algebra (see Beilin & Lust, 1975; p. 187).

(b) *Summary of research findings.* Evidence has been provided to show that some connectives are more frequently used and more easily understood than others. Connectives expressing conjunction (e.g. *and*) are readily understood by very young children (Neimark & Slotnick, 1970; Suppes & Feldman, 1971). Conditional connectives (e.g. *if-then*) and disjunctive connectives (e.g. *or*) appear to be considerably more difficult to understand, the latter giving difficulty until high-school age (Neimark & Slotnick, 1970). These findings are supported in general by Beilin & Lust (1975) and by others (Nitta & Nagano, 1966; Youniss & Furth, 1967; Peel, 1967; Katz & Brent, 1968; Bart, 1974).

These studies provide much insight into the nature and understanding of linguistic connectives; they also reveal how much we still need to know about them. Beilin & Lust (1975) rightly conclude that no comprehensive study has been made of the development of children's understanding of connectives, either in terms of the onset of that understanding (although much can be inferred from child language research, Brown, 1973), or its various stages of development.

There are also large numbers of connectives waiting in the wings ready to be employed. This is to be expected, since the focus of research has been on logical propositional relationships for which there are limited openings for connectives. However, in addition to variants for frequently used connectives (e.g. 'suppose that' for 'if'), there are many connectives that signal relations not typically the focus of research (e.g. connectives expressing relations of time, such as *before, after;* relations of adverseness, such as *however, although,* Pearson, 1976). Strevens (1972) has attempted to categorize such connectives; Gardner (1975) has made a useful compilation, and has begun preliminary explorations of students' understanding of them. Gardner's work reminds us that not all connectives signal the logical propositional relations of conjunction, disjunction and conditionality.

Secondly, these studies lend much support to the hypothesis that a child's level of cognitive development (especially as viewed in a Piagetian framework) is strongly related to his ability to comprehend linguistic connectives. Beilin & Lust (1975) suggest that the development of a child's understanding of linguistic connectives (in particular *and, or* and *not*) closely relate to the development of the child's comprehension of theoretically related truth-functional logical processes (intersection, union, complementation) although they are not related in a one-to-one correspondence (p. 336). This is not entirely surprising, given the choice of linguistic connectives.

However, several questions remain. Firstly, the actual task demands of studies conducted in this area varies enormously. Piaget (1968) used a sentence completion task ("The man fell of the bicycle because ____________ "); Olds (1968), Suppes & Feldman (1971), and Beilin & Lust (1975) asked students to manipulate objects, given verbal instructions ("If you have a circle, you may move it one space"; "Give me the dolls which are girls or give me the dolls which are boys"); Paris (1973) asked students to judge the truth value of statements describing situations pictorially represented. In one sense, these tasks define operationally what is meant by "understanding" a linguistic connective, but how generalizable are these definitions? Are the reasoning processes used by the child in solving the experimenter's tasks the same as those used in the solution of naturally occurring tasks, say in reading a passage from a social studies text? Are there unnatural constraints in the experimental situation not present in natural situations? Is the logic of logical propositional relations a 'natural' logic? Are only logical propositional relations expressed by these linguistic connectives?

A recently published set of papers (Falmagne, 1975) addresses some of these questions. Johnson-Laird (1975), for example, suggests that adults do not use the rules of standard logic in solving syllogistic tasks; rather information from the task undergoes transformations, being also integrated with information from separate sources. Lakoff (1971) presents a similar argument — that the semantic conditions under which propositions are linked are crucial to a person's understanding of them (and, consequently, his understanding of the linguistic connective linking them). Lakoff suggests three influences — (a) the overtness of the propositions — i.e. does the child have previous experience with them, or does he have to derive them by combining pre-suppositions or deduce them through operations on the actual words or meaning? (b) the hierarchy of 'naturalness' — i.e. connectives may be more easily understood when they link universal rather than idiosyncratic propositions; (c) the particular linguistic function of the connective — i.e. symmetric versus asymmetric uses of *and*, which have different semantic and syntactic properties even though the underlying logical operation (i.e. conjunction) is the same (Stahl, 1968). Paris (1973) was aware of this as a limiting factor (p. 290), although neither he nor others appear to have attempted to control for it systematically. The problem of the hierarchy of naturalness is particularly relevant to studies investigating children's understanding of *or* — for example, "Give me the dolls which are not girls or give me the dolls which are not boys" (Beilin & Lust, 1975) is a task that requires a child temporarily to suspend his usual concepts of naturalness.

Staudenmayer (1975) has suggested that we need to distinguish between the process of reasoning used by the child and the resulting evaluations observed. He argues that there are both linguistic (i.e. semantic and syntactic) and non-linguistic (i.e. response bias strategies, interpretation of the task per se, etc.) factors that influence how a subject will understand implicative sentences, but that once these factors have been accounted for, the process of evaluating inferences does in fact follow the rules of logic. One problem here is that it is very difficult to know, from subjects' errors on linguistic connective tasks, where the breakdown has occurred. Is it in the input stage or in the deductive process itself? These considerations may severely limit the generalizability of the studies quoted earlier, although they do appear to confirm the relevance of research on linguistic connectives to the field of cognitive psychology.

Linguistic connectives and reading

(a) *Relevance of the study of connectives to reading research.* Several recent developments in reading research make the study of linguistic connectives particularly relevant to reading. Interest in reading comprehension as a focus of research, particularly beyond the initial grades (1-3) has been spurred by reports of poor reading achievement in the middle and upper grades (National Assessment, 1973; NIE Conference on Reading, 1974) and by the continued efforts of a group of researchers to promote comprehension not decoding as a primary if not sole goal for reading instruction, even in the initial grades (Goodman, 1975;

Smith, 1975). If comprehension is the goal of reading instruction, then one important task for children to learn is to understand complex relations between verbal concepts presented in connected discourse (Smith, 1975). Since linguistic connectives are frequently used to signal these relations (whether semantic, syntactic, logical or a combination of all three), an understanding of their role and function, even implicitly, is clearly indicated.

Secondly, building on the earlier work of Huey (1908) and Thorndike (1917), and more recently incorporating Piaget's theoretical framework for cognitive development, several researchers have promoted a view of reading as reasoning (Thorndike, 1973; Stauffer, 1975) or as concept development (Henry, 1974), not as easily fractionated into separable components as was previously supposed (Davis, 1968; Smith & Barrett, 1974).

It might follow that in reading for understanding, experience and competence in understanding the logical structure of connected discourse is important, and therefore learning the process of deductive thinking might be an essential component of learning to read/think. In this regard, understanding the linguistic connective's role and function is particularly relevant. Henry (1974) recognizes this, and presents suggestions for learning linguistic connectives in the context of such aspects of concept development as resolving contradictions (p. 35ff).

However, very little attention has been paid to the connective in the reading research literature, although it does merit passing reference by Huey (1908), Thorndike (1917), Flesch (1947), Smith (1963), and others.

Another glancing blow is struck by Pearson (1976), who provides evidence that sentences made 'easier' to read by the removal (amongst other things) of linguistic connectives are not always in fact more easily understood by children. However, Pearson's orientation and purpose was not to provide direct evidence of children's knowledge and use of the connectives.

(b) *Brief review of studies on linguistic connectives and reading.* The review is brief, principally since so little has been done on linguistic connectives from the reading perspective. Two recent studies, however, are relevant. Robertson (1968) examined fourth, fifth and sixth graders' reading comprehension of connectives; Gardner (1975) examined seventh to tenth graders' understanding of them. Both studies, completely independently (one in Canada, the other in Australia), used essentially the same technique. Robertson examined basal readers to determine which connectives to test, finding, in the process, that more than one third of all sentences sampled contained connectives, and three-quarters of these contained one connective, the remainder at least two. Gardner examined secondary science materials, noting frequencies of all logical connectives. Items for tests in both cases were similar; Robertson used sentence completion items in which the stem contained the connective, with multiple choice predicates. Gardner used this approach, also a modified cloze in which the connective was omitted but a list of alternatives was provided.

Robertson found that children's understanding of the connectives improved from 57% in grade 4 to 66% in grade 5 to 75% in grade six; and that a significant relationship exists between comprehension of connectives and sex, mental age (verbal and nonverbal), and ability in listening, reading and written language. Chronological age and socioeconomic status appeared unrelated.[1]

Gardner's data so far are from pilot studies, and his analyses were primarily aimed at test development (including comparisons between sentence completion and modified cloze), thus no firm conclusions may yet be made. However, he does speculate that his sentence completion items "measure a highly internally consistent cognitive skill . . . (possibly) the ability to process separate simple propositions and re-assemble them into complex propositions" (p. 14).

Two conclusions at least may be drawn from these studies. Firstly, there are many connectives in the materials used by children beyond the initial grades, both in general reading matter and in specialized content areas such as science. This may be confirmed by reference to Carroll's *Word Frequency Book,* in which for example connectives occupy 34 out of the first 100 most frequently appearing words (Carroll, 1971). However, no count appears to have been made of the frequency and type of concepts related by these connectives. Secondly, it would appear that a child's understanding of connectives (as defined and measured in the Robertson/Gardner studies) does correlate with measures of general reading ability and achievement, and measures of intellectual functioning; although the measures used by Robertson and by Mahoney & Domaracki (1976) were too global to permit much inference.

(c) *Implications for further study.* There needs to be some resolution of the two approaches towards studying children's knowledge of linguistic connectives. The approach that sees connectives as signals for underlying logical propositional relations (Paris, 1973; Beilin & Lust, 1975) derives its test items from a logico-grammatical base that is internally consistent but not necessarily consonant with natural language processing. Further, it may not adequately account for linguistic (i.e. semantic and syntactic) relationships between propositions (Lakoff, 1971) nor for the influence of non-linguistic factors (i.e. response bias, etc., Staudenmayer, 1975). While this approach considerably enhances our understanding of the logical base of language, its focus is narrow and consequently its application to empirical language processing tasks may be of limited utility.

The approach that sees connectives in terms of their frequency of occurrence in natural language (or in school-related materials) derives its test items from an examination of natural written discourse that is internally consistent, but it has failed to fully appreciate the nature of the underlying logical propositional relations signalled by the connective. This approach, too, may not have adequately considered the linguistic and non-linguistic factors as described above. Certainly, the tasks required of subjects in the two approaches appear to be markedly dissimilar, and it has been shown that task demands influence results considerably. Some experimentation in the reading field with alternative approaches to tasks is clearly indicated.

Secondly, the work already begun by Strevens (1972) and continued by Gardner (1975) categorizing logical connectives needs to be developed further.

---

[1] An unpublished study, using Gardner's approach with 9th graders showed positive relationships between comprehension and reading achievement. There were no sex differences, but marked differences between sentence completion and modified cloze items (Mahoney & Domaracki, 1976).

Studies which have focused primarily on connectives that signal logical propositional relations (e.g. conjunction, disjunction, conditionality) should probably be extended to include more 'natural' relations such as *time* (e.g. before, after), *adverseness* (however, although), *purposiveness* (so that, in order to), *elaboration* (that is, in other words) and so on (Pearson, 1976). These may be useful categories into which connectives found in running text might be placed.

Thirdly, an examination of the role and function of connectives in reading comprehension might usefully be made through the investigation of good and poor readers' understanding of them. If an important purpose of reading is to gain access to underlying relations (e.g. cause and effect, conditionality, contrast, etc.), then the study of the use made of the connective by readers in gaining this access is extremely relevant. It is likely that good readers do not use connectives in the same way as poor readers — for example, they may have more experience of the topics being related; or more highly developed deductive reasoning processes; more efficient attentional mechanisms. Poor readers may be hampered by deficiencies in any or all of these. Such investigations would not only enlarge the knowledge base of reading processes but also yield information directly useful for classroom practices.

REFERENCES

Bart, W. M. Adolescents' understanding of logical connectives. *Psychological Reports*, 1974, *35*, 41-42.

Beilin, H., Lust, B., Sack, H., & Natt, H. *Studies in the cognitive basis of language development.* New York: Academic Press, 1975.

Brown, R. *A first language.* Cambridge: Harvard University Press, 1973.

Carroll, J. B., Davies, P., & Richman, B. *Word Frequency Book.* Boston: Houghton Mifflin, 1971.

Davis, F. B. Research in comprehension in reading. *Reading Research Quarterly*, 1968, *3*, 499-545.

Falmagne, R. J. (Ed.). *Reasoning: representation and process in children and adults.* Hillsdale, N.J.: Erlbaum, 1975.

Flesch, R. *The art of plain talk.* New York: Harper, 1947.

Gardner, P. L. *Logical connectives in science: a preliminary report.* Paper presented to the 6th Annual Conference of the Australian Science Education Research Association, May, 1975.

Goodman, Y. Reading comprehension: A redundant phrase. In N. A. Johnson (Ed.), *Issues in language.* New York: Winthrop, 1975.

Henry, G. H. *Teaching reading as concept development.* Newark, Delaware: IRA, 1974.

Huey, E. B. *The psychology and pedagogy of reading.* Cambridge, Mass.: MIT Press, 1968. (Originally published, 1908).

Johnson-Laird, P. N. Models of deduction. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults.* Hillsdale, N.J.: Erlbaum, 1975.

Katz, E. & Brent, S. Understanding connectives. *Journal of Verbal Learning and Verbal Behavior*, 1968, *7*, 501-509.

Lakoff, R. If's, and's, and but's about conjunction. In C. J. Fillmore & D. T. Langendoen (Eds.), *Studies in linguistic semantics.* New York: Holt, 1971.

Mahoney, P. & Domaracki, J. *Adolescents' understanding of logical connectives and reading comprehension.* Unpublished paper, State University of New York at Albany, 1976.

National Assessment. *Reading: Themes 1-8.* Denver, Colorado: Education Commission of the States, 1973.

National Institute on Education. *Conference Proceedings on Reading.* Washington: NIE, 1974.

Neimark, E. D. Development of comprehension of logical connectives: Understanding of "or" *Psychonomic Science*, 1970, *21*, 217-219.

Neimark, E. D. & Slotnick, N. S. Development of the understanding of logical connectives. *Journal of Educational Psychology*, 1970, *61*, 451-460.

Nitta, N. & Nagano, S. Basic logical operations and their verbal expressions. *Research Bulletin of the National Institute for Educational Research*, 1966, *7*.

Olds, H. F. *An experimental study of syntactical factors influencing children's comprehension of certain complex relationships.* Final Report. Cambridge, Mass.: Center for Research and Development in Educational Differences, 1968.

Paris, S. G. Comprehension of language connectives and propositional logical relationships. *Journal of Experimental Child Psychology,* 1973, *16,* 278-291.

Pearson, P. D. The effects of grammatical complexity on children's comprehension, recall and conception of certain semantic relations. In H. Singer & R. B. Ruddell (Eds.), *Theoretical models and processes of reading.* Newark, Delaware: IRA, 1976.

Peel, E. A. A method for investigating children's understanding of certain logical connectives used in binary propositional thinking. *British Journal of Mathematical and Statistical Psychology,* 1967, *20,* 81-92.

Piaget, J. *Judgment and reasoning in the child.* New Jersey: Littlefield Adams, 1968.

Robertson, J. E. Pupil understanding of connectives in reading. *Reading Research Quarterly,* 1968, *3,* 387-417.

Shapiro, B. J. & O'Brien, T. C. Logical thinking in children ages 6 through 13. *Child Development,* 1970, *41,* 823-829.

Smith, D. V. Growth and sequence of language. In H. A. Robinson (Ed.), *Reading and the language arts.* Supplementary Educational Monographs, 1963, *93.*

Smith, F. *Comprehension and learning.* New York: Holt, 1975.

Smith, R. J. & Barrett, T. C. *Teaching reading in the middle grades.* Reading, Mass.: Addison-Wesley, 1974.

Staal, J. F. 'And'. *Journal of Linguistics,* 1968, *4,* 79-81.

Staudenmayer, H. Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults.* Hillsdale, N.J.: Erlbaum, 1975.

Stauffer, R. G. *Directing the reading-thinking process.* New York: Harper & Row, 1975.

Strevens, P. *Technical, technological and scientific English.* Paper presented to the Conference of the International Association of Teachers of English as a Foreign Language, London, 1972.

Suppes, P. & Feldman, S. Young children's comprehension of logical connectives. *Journal of Experimental Child Psychology,* 1971, *12,* 304-317.

Thorndike, E. L. Reading as reasoning: A study of mistakes in paragraph reading. *Journal of Educational Psychology,* 1917, *8,* 323-332.

Thorndike, R. L. Reading as reasoning. *Reading Research Quarterly,* 1973, *9,* 137-147.

Vygotsky, L. S. *Thought and language.* Cambridge: MIT Press, 1962.

Youniss, J. & Furth, H. G. The role of language and experience on the use of logical symbols. *British Journal of Psychology,* 1967, *58,* 435-441.

309